United States
Environmental Protection
Agency

Workshop Report

Developing a Comprehensive Long-Term Research Strategy to Support Determination of Protective Ballast Water Discharge Standards

U.S. Environmental Protection Agency
Ecological Exposure Research Division
National Exposure Research Laboratory
Research Triangle Park, NC 27711

Final revised draft
June 30, 2013

EXECUTIVE SUMMARY

Unmanaged ballast water discharges have resulted in the introduction of aquatic nuisance species (ANS) that have led to severe ecological degradation and billions of dollars in damages. Until recently, ballast water management has focused on utilizing ballast water exchange or saltwater flushing. However, due to the limitations of these techniques, the regulatory focus has shifted toward ballast water performance standards that limit the number of living organisms discharged into receiving waters.

EPA and the U.S. Coast Guard sought the advice of the National Academy of Sciences National Research Council (NRC) to do the following: 1) evaluate the state of the science of various approaches to assess the of ANS establishment associated with ballast water, 2) recommend how regulatory agencies can use these approaches to best inform risk management decisions on allowable concentrations of ANS in discharged ballast water to protect against ANS establishment, and 3) evaluate the risk of successful establishment of new ANS associated with ballast water discharge limits used or suggested by international and domestic regulatory agencies. The NRC found a profound lack of information that inhibited the development and validation of models, preventing the determination of risk of ANS establishment under existing discharge limits. The goal of this document is to further advance a cohesive binational strategy (United States and Canada) to fill some of these critical data gaps. In order to meet this goal, a workshop was convened in September 2012, including participants with expertise in ballast water risk assessment, aquatic nuisance species biology, theoretical population biology, and the design of experimental and statistical approaches to understanding relationships between propagule supply and establishment likelihood. This report focuses on the following issues addressed during that workshop:

1. Models
2. Existing Datasets
3. Experimental Studies
4. Shipboard Surveillance
5. Designing a Long-Term Port Surveillance Program
6. Selection of Study Targets
7. Incorporation of Genetic Tools
8. Logistical Considerations, Coordination, and Benefits of Proposed Research

Models

This report discusses potential useful models that can be parameterized with experimental data and used to characterize population establishment in relation to propagule pressure from ballast water discharges. The aim of the workshop was not to identify specific models to pursue, but rather to assess data needs of the most promising models and recommend a course of action toward filling those needs. The report discusses the strengths and weaknesses of three probabilistic and dynamic demographic models, including population viability analysis, stock-recruitment models, and epidemiological models. Also discussed are the strengths and

weaknesses of using a per capita invasion probability (PCIP) statistical model. Mechanistic models will ultimately be the most useful for management because they enable the evaluation of different management strategies under future scenarios. However, fitting these models may require more data than are available, in which case hybrid mechanistic-descriptive models can be used until more data are collected. Identifying, collecting, and meta-analyzing existing datasets compose a promising and cost-effective approach toward filling the critical data gaps.

Common data requirements allow different models to use the same data to test for agreement. These requirements include instantaneous growth rate and variation along with biotic and abiotic variables that affect population growth. The primary challenge will be to convert these laboratory and mesocosm results into a vessel-scale management threshold, which requires determining how the actual invasion risk from many species in many ships across many locations differs from the predicted risk of one species in a controlled experiment.

Existing Datasets

A recent literature review looked at 66 published references reporting on animal and plant species identified in ballast water or ballast sediment during sampling efforts since 1973. Fifty-six percent of these sampling efforts focused on North American ports. The studies varied by sampling method, number of samples taken, number of ships sampled, spatial resolution, temporal resolution, and size of target organism. Most publications described the sampling methods and preservation and analysis of samples. However, few commented on whether samples had been vouchered and stored for future use and the raw data are generally not publically available. Many of these publications do give detailed lists at genus and species level of taxa observed.

Datasets such as these provide a starting point for estimating propagule pressure (PP) and understanding uncertainty surrounding such estimates. Unfortunately, previous studies suggest that our ability to confidently estimate propagule pressure from existing data may be limited, as propagule supply varies dramatically with a large number of factors. Nevertheless, this report recommends developing a coordinated database of raw data on ballast water samples, with the aim of incorporating existing data and providing a standardized repository for future data. Such a database will facilitate formal meta-analysis of available ballast water sampling data, which could yield insights into the factors driving variation in propagule pressure to improve existing models. This could also identify appropriate sampling strategies for collecting additional data and inform selection of future models. A database of raw data could be adopted as a central repository for similar data collected in the future, providing a framework for standardizing reporting and an opportunity to effectively leverage independent efforts. Such a dataset could have significant value in identifying temporal patters of propagule supply related to shifting regulatory climates.

Experimental Studies

Experimental studies provide a cost- and time-efficient way of obtaining critical data, as well as inform the theory of invasion dynamics, and they parameterize and validate risk-release models. Experiments allow control of the propagule pressure and exposure of target communities to many different conditions while measuring the outcome. These data may be used to

parameterize or validate models, which could be used for setting conservative discharge standards in order to limit invasion risk.

Experiments may be conducted at the benchtop-, mesocosm-, or field-scale. Whole-ecosystem field studies are often infeasible because they are difficult to control, hard to replicate, expensive, and run the risk of introducing ANS to uninfested waters. Benchtop-scale experiments provide the most control over environmental factors. Mesocosm-scale experiments moderate the risks and benefits of the scale extrema. However, the accuracy of predictive models, and their value in regulatory contexts, depends on whether scaling from bench- and meocosm-level to ecosystem level is handled properly.

Experimental studies should focus on taxa that represent worst-case scenarios for invasion. Given the likely need to conduct such studies at the mesocosm level and the limited availability and relatively high cost of such experiments, resources are best applied to understand the factors underlying the risk-release relationship for single species, with the acknowledgement that single-species systems are simplified models of ecosystems. Experimental systems must be able to support establishment of target species at some relevant inoculum level, within an experimentally tractable time frame, and data collection should focus on temporal sampling of population density as well as experimental endpoints (e.g., success/failure of establishment) so that experiments can support development of both statistical models of establishment success *and* dynamic demographic population models. Whenever possible, experiments should be conducted at multiple scales (e.g., benchtop and mesocosm). Confidence in reliability of experimentally derived parameters will be increased if multiple facilities can coordinate and replicate experimental studies, which would entail adoption of standardized experimental protocols designed to minimize potential biases between sites and to ensure that variation in results can be attributed to relevant biological differences.

Shipboard Surveillance

While it is not feasible to expect to measure all sources and densities of propagules or identify all introduced species, limited, strategic shipboard sampling coupled with appropriate models can be used to make valid statistical inferences about the distribution of propagule densities per vessel and the overall load to the ecosystem. In designing shipboard surveillance programs, it is important to consider both the vessel type and sampling method (e.g., in-line vs. ballast tank sampling). Development of a standardized approach and widely accepted ship classification scheme is a critical step in designing optimal sampling approaches. The report notes that due to changes in regulated behavior, the risk from coastal domestic merchant vessels is higher than from international merchant vessels. With respect to sampling techniques, the report contrasts the strengths and weaknesses of in-tank sampling of ballast tanks with the advantages and disadvantages of the increasingly common in-line sampling.

The primary aim of shipboard surveillance as described here is to sample enough ships and account for enough proportion of the variance across multiple factors known to influence propagule supply to characterize the probability distribution of propagule densities per vessel and the overall delivery to the entire ecosystem. This report recommends long-term intensive (likely >100 vessels/year) sampling of three to six target ports, ideally the same as those selected for surveillance of recipient environments. Funding limitations favor more thorough sampling at

fewer sites over less intensive sampling at more sites, as it is critically important to develop a comprehensive picture of propagule supply patterns to at least one of the selected target ports during implementation of the recently established discharge standards. In-line sampling approaches can be readily standardized and are likely to be implemented more broadly in the future; therefore, they should be adopted as the standard approach for sampling. An approach to assess the optimal stratification of sampling within ships, between ships within ports, and between ships in different ports, should be adopted, taking into account the estimated increased accuracy of more accurate propagule supply estimates using more sampling units and the expected costs to add them. Importantly, analyzing existing data sets (some collected over >10 years) will add insight by provided additional information on the importance of different factors in determining variation in propagule supply. Overall estimated costs associated with this intensive ballast water surveillance effort could be as much as $1 million per target port over a 10-year period.

Designing a Long-Term Port Surveillance Program

Design of a coordinated recipient system surveillance program must satisfy a number of criteria. It must effectively detect rare taxa; it must generate data that are comparable across surveys within the same system and across multiple target systems; it must be statistically robust; and it must achieve the greatest possible cost efficiency. Substantial guidance on systems in Australia, New Zealand, and the United States already exists on designing effective surveillance efforts aimed at determining non-native diversity and detecting novel invasions and can be used to inform the design of appropriate surveillance approaches.

This report recommends aggressively pursuing design of surveillance strategies for three to six target ports. Concentrating effort at fewer sites will increase cost efficiency, and comprehensive understanding of the relationship between propagule supply and establishment at even one site will vastly improve our ability to assess various models of the risk-release relationship. At each of the target ports, surveillance should include both a targeted assessment of presence/absence and distribution of a small set of taxa and non-targeted overall assessment of biodiversity. Most of the investment in long-term surveillance should examine total diversity and focusing on sampling key habitats and key taxonomic groups that commonly use ballast water as a vector (in an effort to distinguish among the effects of organisms delivered by ballast water vs. hull fouling). Efforts to assess total diversity in samples should include applying traditional approaches to taxonomic identification, and novel methods of taxonomic identification (e.g., genetic tools) that allow simultaneous, cost-effective processing of increasingly higher numbers of individuals. Statistical methods should be adopted that enable estimation of richness of non-native species, even when not detected. If appropriate sampling design can make inferences of species diversity sufficiently statistically robust, total numbers of established non-native species may possibly be estimated, even if their identities are unknown.

The report recommends designs based on a combination of the Australian Centre for Research on Introduced Marine Pests (CRIMP) protocols and passive sampling methods with incorporation of genetic analysis, and advocates for surveillance design to be adaptive. This approach will be facilitated by tight coordination within and among surveillance efforts and established protocols for formal project review and oversight throughout the lifetime of the program. Detection likelihoods should be frequently assessed and encounter functions should be

estimated using rarefaction (species accumulation curves) and other established methods. Sampling strategies should be refined as necessary to account for new information on recipient system environmental or biotic parameters. The adaptive design approach will be particularly important for incorporating new technologies (e.g., genetic methods) that have yet to reach full maturity in the context of early detection and monitoring. It is reasonable to suppose that a recipient system surveillance program for a single port over 10 years could require $5 million, resulting in an overall cost of surveys for three to six ports in the $15- to $30-million dollar range.

Selection of Study Targets

This report recommends focusing on taxa that present worst-case scenarios for invasion, in contrast to selecting "representative" invaders. This recommendation is based primarily on pragmatic considerations related to the feasibility of the research effort and the likelihood of developing models that are most useful to managers. Other efforts, particularly non-target approaches, will expand understanding of the risk-release relationship beyond organisms that have been considered, *a priori*, high risk for establishment, which is critically important not only to prevent model bias, but also to guard against the possibility that those *a priori* assessments of risk fail to accurately capture the true likelihood of establishment for chosen target species.

Factors that determine appropriate targets for intensive study include: biological characteristics predisposing taxa to invasiveness (capacity for asexual reproduction, high fecundity, trophic or habitat generalist), ease of sampling and identification (including the existence of developed molecular probes to facilitate use of genetic methods), ability to prevent accidental release or quarantine of experimental organisms, and ability to confidently link non-native establishment to ballast water. Selecting multiple study taxa is desirable. Also important will be selecting at least some study taxa that will be adopted for both experimental and descriptive ("surveillance") approaches.

Due to cost and time considerations, comprehensively assessing the non-native biodiversity present in North American coastal systems is beyond the scope of any attainable research effort. Likewise, capturing statistically satisfying variation among ports (requiring sampling of a large number of ports) may well be beyond the scope of the described research effort. Intensive surveillance will be possible only in a small subset of existing recipient ports, and appropriate port selection is therefore critical. Selected ports should have well-described invasion histories and historical patterns of ballast water exchange; have well-described baseline biodiversity (including referenced genetic data); have substantial variation in sources of ballast water; and be capable of providing substantial logistical support. Examples of such ports in North America include Chesapeake Bay, San Francisco Bay, Tampa Bay, and Duluth/Superior Harbor.

Incorporation of Genetic Tools

This report recommends incorporating genetic detection methods into both ship and port surveillance efforts to take advantage of potential efficiencies and minimize cost per unit effort. Genetic methods may be used to replace or augment monitoring based on traditional morphological identification. The report recommends developing targeted DNA-based detection

methods based on quantitative polymerase chain reaction (qPCR) approaches for any target species selected for surveillance coupled with non-targeted community profiling using next-generation sequencing (NGS) methods, with an emphasis on the latter. It also discusses the advantages and disadvantages associated with specific tools. In particular, while genetic tools potentially bring numerous improvements over existing methods, it is important to note that genetic technologies for surveillance are still in their infancy and numerous technical hurdles remain to be overcome (e.g. challenges associated with processing large sample volumes and generating quantitative abundance estimates). Nevertheless, the increased cost of incorporating such tools, particularly in early stages of surveillance when these technologies are still in development, should be well worth the increased sensitivity of detection and taxonomic resolution. More importantly, as the technology advances and confidence increases in the inferences drawn from genetic data, considerable future costs can be saved by relieving some of the burden associated with traditional morphological identifications. Estimated costs are considered and have been incorporated into overall cost estimates for surveillance efforts.

Logistical Considerations, Coordination, and Benefits of Proposed Research

Although considerable effort has already been invested into research to understand the risk-release relationship, these studies have not been coordinated. This report provides guidance for developing an ambitious research program that seeks to efficiently leverage financial and human resources across multiple projects to generate data most salient for informing risk release models. This frees experimental and descriptive approaches to remain geographically independent, and also allows experimental studies to be funded independently, which may be an important overall design consideration. In contrast, ship and port surveillance efforts absolutely require very tight coordination. Benefits of such coordination of ship and port surveillance include: cost efficiency, standardization, managerial oversight facilitating adaptive research planning, coordinated funding, and visibility and public outreach.

While the primary aim of the proposed research is to inform our understanding of the ballast water risk-release relationship, the benefits of the broad, coordinated research effort described in this report would range far beyond the relatively narrow interest of parameterizing models relating propagule supply to establishment risk. This ancillary benefits include better general understanding of how propagule supply relates to ANS establishment likelihood, which may ultimately be relevant to understand risk release relationships for other vectors of introduction; increased general knowledge of the mechanisms and patterns of colonization success in aquatic systems; establishment of a program in target systems for registering changes in propagule supply and invasion rate associated with management and policy changes; and valuable lessons in designing effective early ANS detection and monitoring at the level of large ports of entry, with additional relevance to design of more general biodiversity surveillance efforts. Consideration of these ancillary benefits is critically important when assessing the likely return on investment associated with substantial expenditure of public funds, particularly in challenging fiscal climates such as the one we currently face. This report recognizes these fiscal challenges, and further provides recommendations for potential downscaling and prioritization of the larger research effort.

WORKSHOP PARTICIPANTS AND CONTRIBUTORS

Name	Affiliation
Ryan Albert	US Environmental Protection Agency
Sarah Bailey	Fisheries and Oceans Canada
Peg Brady	NOAA
Allegra Cangelosi	Northeast-Midwest Institute
Robin Danesi	US Environmental Protection Agency
John Darling	US Environmental Protection Agency
Meghana Desai	Northeast-Midwest Institute
Nancy Elder	US Geological Survey
Alan Ellsworth	US National Park Service
Rich Everett	US Coast Guard
Jon Geller	Moss Landing Marine Laboratory
Erin Grey	University of Notre Dame
Colin Henein	Transport Canada
Mike Hoff	US Fish and Wildlife Service
Chris Jerde	University of Notre Dame
Marten Koops	Fisheries and Oceans Canada
Brian Leung	McGill University
David Lodge	University of Notre Dame
Hugh MacIsaac	University of Windsor
Nick Mandrak	Fisheries and Oceans Canada
Christine McLaughlin	US Environmental Protection Agency
Amanda Newsom	California State Lands Commission
Susan Pasko	NOAA
Gregory Ruiz	Smithsonian Environmental Research Center
Doug Schneider	World Shipping Council
Scott Smith	US Geological Survey
Matthew TenEyck	University of Wisconsin Superior
Chris Wiley	Transport Canada

ACRONYM LIST

Acronym	Definition
ABWMAC	Australian Ballast Water Management Advisory Committee
ANS	Aquatic Nuisance Species
CAISN	Canadian Aquatic Invasive Species Network
CRIMP	Center for Research on Introduced Marine Species
CSIRO	Commonwealth Scientific and Industrial Research Organization
CWA	Clean Water Act
GLSLR	Great Lakes St. Lawrence Region
HIS	Habitat Suitability Index
IMO	International Maritime Organization
NAMPCA	Nonindigenous Aquatic Nuisance Prevention and Control Act
NISA	National Invasive Species Act
NGS	Next Generation Sequencing
NIH	Nonindigenous Species
NRC	National Academy of Sciences National Research Council
PCIP	Per Capita Invasion Probability
PCR	Polymerase Chain Reaction
PP	Propagule Pressure
PVA	Population Variability Analysis
RAG	Research Advisory Group
SEIR Model Subgroup	Susceptible-Exposed-Infective-Recovered Model Subgroup
S-I Models	Susceptible-Infective Models
SIR Model Subgroup	Susceptible-Infective-Recovered Model Subgroup
S-R Models	Stock-Recruit Models
VGP	Vessel General Permit

CONTENTS

Page

EXECUTIVE SUMMARY ..... **II**

WORKSHOP PARTICIPANTS AND CONTRIBUTORS ..... **V**

ACRONYM LIST ..... **VI**

1. INTRODUCTION **1**

2. WHICH MODELS ARE AVAILABLE? **5**
2.1 Probabilistic and Dynamic Demographic Models 6
2.1.1 Population Viability Analysis 6
2.1.2 Stock-Recruit Models 7
2.1.3 Epidemiological Models 8
2.2 Statistical Models: Per Capita Invasion Probability (PCIP) 9
2.3 Common Data Needs 11

3. LEVERAGING EXISTING DATASETS **13**
3.1 Overview 13
3.2 Estimating Propagule Pressure from Descriptive Datasets 16
3.3 Estimating Modifiers (ε) from Descriptive Datasets 18
3.4 Estimating Establishment Probability (P_E) from Descriptive Datasets 19
3.5 Recommendations for Analysis of Existing Datasets 19

4. EXPERIMENTAL STUDIES **21**
4.1 Aims of Experimental Studies 21
4.2 Past/Ongoing Experimental Studies 22
4.3 Issues of Scale and Extrapolation 24
4.4 Design Recommendations 26

5. DEVELOPING A SHIPBOARD SURVEILLANCE STRATEGY **29**
5.1 How Much Sampling is Needed? 29
5.2 Classification of Vessels/Pathways to Inform Choice of Surveillance Targets 30
5.3 Comparison of Ballast Sampling Approaches 32
5.4 Recommendations for Shipboard Surveillance 33
5.5 Cost Considerations 35

6. DESIGNING A LONG-TERM PORT SURVEILLANCE PROGRAM **37**
6.1 General Considerations 37
6.2 Design of Sampling and Analytical Approaches 38
6.2.1 Examples of National-Scale Standardized Surveillance Efforts 38
6.2.2 Recommendations for Long-Term Port Surveillance 42
6.3 Cost Considerations 44

7. SELECTION OF STUDY TARGETS **46**
7.1 Selection of Study Taxa 46
7.2 Selection of Target Ports 47

CONTENTS (Continued)

Page

8. INCORPORATION OF GENETIC TOOLS 49
8.1 Background 49
8.2 Targeted Detection Tools 50
8.3 Community Profiling Tools 51
8.4 Cost Considerations 53

9. LOGISTICAL CONSIDERATIONS, COORDINATION, AND BENEFITS OF THE PROPOSED RESEARCH 56
9.1 Building a Coordinated Research Effort 56
9.2 Ancillary Benefits of the Research Program 59
9.3 Possibilities for Down-Scaling and Prioritization 61

10. REFERENCES 63

11. APPENDICES 80
11.1 Abbreviated Summary of Published Reports on Ballast Water Sampling 80
11.2 Brief Survey of Available Genetic Technologies 86

LIST OF TABLES

Page

Table 2-1. Data requirements of different types of risk-release models (categories derived from Wonham et al. (2013)). 12

Table 6-1. Costs associated with nuisance species surveillance programs in terrestrial and aquatic systems 45

Table 11-1. Comparison of technologies available for DNA-based detection and monitoring. 91

LIST OF FIGURES

Page

Figure 3-1. Number of publications (y axis) per 5-year period between 1975 and 2012 (x axis) 14

Figure 3-2. Locations of ballast water survey studies for publications between 1975 and 2012..... 15

Figure 3-3. Frequency histogram showing the distribution of publications that sampled a given number of vessels..... 15

Figure 9-1. Proposed Timeline for Research 59

1. INTRODUCTION

Unmanaged ballast water discharges have resulted in the introduction of aquatic nuisance species (ANS) that have led to severe ecological degradation and billions of dollars in economic damages (Mills et al. 1993, Ruiz et al. 2000, Wonham and Carlton 2005, Lovell and Drake 2009, Carlton et al. 2011, 77 Federal Register 17254 2012). Until recently, ballast water management focused on utilizing ballast water exchange or saltwater flushing. Ballast water exchange involves replacing coastal water in ballast tanks by mid-ocean water (Gregg et al. 2009). Saltwater flushing involves pumping ocean water into empty ballast tanks, which often have small volumes of residual water and sediment, where it is mixed before pumping it back out mid-ocean (Ruiz and Reid 2007, Briski et al. 2010). Ballast water exchange began to be implemented as a voluntary guideline for the Laurentian Great Lakes in 1989 and was later adopted as a voluntary guideline by the International Maritime Organization (IMO) in 1991. It became mandatory in the United States, first in the Great Lakes under the Nonindigenous Aquatic Nuisance Prevention and Control Act (NANPCA) and then nationally under the Nonindigenous Aquatic Nuisance Prevention and Control Act/National Invasive Species Act (NISA). Canada made ballast water exchange and saltwater flushing mandatory under the Canada Shipping Act in 2006. Saltwater flushing and ballast water exchange are also mandated in the United States by the St. Lawrence Seaway Development Corporation and the U.S. Environmental Protection Agency (EPA). The cooperative research and regulatory development that has led to these regulatory developments in Canada and the United States has continued through coordinated and successful enforcement efforts to reduce invasion risks, particularly in shared waters such as the Great Lakes-St. Lawrence Seaway system (Bailey et al. 2011).

In recent years, awareness of limitations of ballast water exchange and salt-water flushing procedures led to a desire to move toward ballast water performance standards that limit the number of viable organisms discharged into receiving waters. Establishing regulatory limits that substantially reduce the risks from ballast water discharges has been a goal of the U.S. Government, many in the international regulatory community, and other stakeholders for more than 20 years. At the international level, ballast water discharges are primarily addressed by the IMO through the *International Convention for the Control and Management of Ships' Ballast Water and Sediments, 2004*, which has been ratified by 38 parties, including Canada, but has not yet entered into force; Transport Canada is working to implement the convention in regulations in consultation with scientific advisors at Fisheries and Oceans Canada. In the United States, ballast water is primarily regulated by two federal agencies. The United States Coast Guard (Coast Guard) regulates ballast water under the authority of NANPCA/NISA, as amended (16 U.S.C. §§ 4701 et seq.). Due to a 2005 court decision, EPA regulates ballast water under Section 402 of the Federal Water Pollution Control Act, commonly referred to as the Clean Water Act or "CWA" (33 U.S.C. §§ 1251 et seq.). Furthermore, several U.S. states have supplemental ballast water management requirements. For additional discussion of international, U.S. domestic, and U.S. state ballast water regulatory framework (including its history), please see Albert et al. (2013), Gregg et al. (2009), or Gollasch et al. (2007).

The IMO, Canada, Coast Guard, EPA, and various U.S. states have adopted numeric concentration-based ballast water standards as a way to manage ballast water discharges. The IMO D-2 standards express limits as three size groups of organisms (IMO 2004b):

- Organisms greater than or equal to 50 µm (to address zooplankton and macrofauna). The limit is less than 10 living organisms per cubic meter.

- Organisms greater than or equal to 10 but less than 50 µm (to address phytoplankton). The limit is less than 10 living organisms per cubic milliliter.

- Specified microorganisms (to address indicator organisms and pathogens). The limits are for *Vibrio cholerae* < 1 CFU per 100 ml (or <1 CFU gram [wet weight] of zooplankton samples); *E. coli* < 250 CFU per 100 ml; and Intestinal enterococci < 100 CFU per 100 ml.

In March 2012, in its ballast water discharge standard rulemaking, the Coast Guard finalized a ballast water discharge standard was essentially the same as the IMO that, like the IMO standard, must be met by ship owners on a rolling implementation schedule (77 Federal Register 17254 2012). In March 2013, EPA finalized the Vessel General Permit (VGP), which contains the same numeric limits and implementation schedule (US EPA 2013). Both the Coast Guard and EPA express those limits as instantaneous maximum limits (77 Federal Register 17254 2012, US EPA 2013). The risk of successful invasions increases with propagule pressure (Lee et al. 2013), and the limits are designed to reduce the number of living organisms discharged, subsequently reducing propagule pressure. However, while some, including Carlton et al. (2011), argue that the IMO D-2 standards discussed above are, at minimum, a good first step toward significantly reducing new invasions, others argue that the limits are not sufficiently protective of water quality (e.g., Great Lakes Alliance 2012, Martens 2012, NRDC 2012, NWF 2012).

When developing the 2013 VGP, EPA, in cooperation with the Coast Guard, sought the advice of the National Academy of Sciences National Research Council (NRC) to identify and apply relevant risk assessment methodologies (Hanlon et al. 2010, Carlton et al. 2011). EPA recognized the need for a robust, scientifically based risk assessment methodology(ies) to inform development of ballast water limits that protect U.S. water quality (US EPA 2013). EPA and Coast Guard researchers prepared a background white paper for that panel that presented several possible risk assessment methodologies and their strengths and weaknesses (Lee et al. 2010) and asked the NRC to:

1. Evaluate the state of the science of various approaches that assess the risk of aquatic nonindigenous species establishment given certain concentrations of living organisms in ballast water discharges.

2. Recommend how regulatory agencies can use these approaches to help determine the allowable concentrations of living organisms in discharged ballast water to safeguard against the establishment of new aquatic nonindigenous species and to protect and preserve existing indigenous populations of fish, shellfish, and wildlife and other beneficial uses of the nation’s waters.

3. Evaluate the risk of successful establishment of new aquatic nonindigenous species associated with a variety of ballast water discharge limits that have been

used or suggested by the international community and/or domestic regulatory agencies.

In its June 2011 report, the binational NRC panel found that, among other things, it is not currently possible to accurately quantify or estimate the risk of ANS or invasive species establishment with any precision. NRC stated that there is "a profound lack of data and information to develop and validate models," and "it was not possible with any certainty to determine the risk of nonindigenous species establishment under existing discharge limits" (Carlton et al. 2011). Based in large part on this advice, EPA found that this *profound lack of data* prevented the Agency from calculating a numeric limit designed to protect water quality (i.e., establishing a numeric concentration-based numeric limit based on an invasion risk assessment) (US EPA 2013). However, as part of its report, the NRC suggested a path forward that involves filling in data gaps needed to develop and validate models, which would help scientists and regulatory agencies better estimate risk in future assessments (Carlton et al. 2011).

As part of the recommended path forward, the NRC suggested selecting a model or models as the foundation for data gathering, then collecting both experimental and field-based data needed to calibrate and, to the extent feasible, validate those models. Generating experimental data would provide results over the short term and would allow researchers and/or regulatory agencies to evaluate the risk-release relationship. Generating field-based data would provide real-world validation and parameterization; however, longer time scales are needed to produce those data.

NRC suggested experimental designs using large-scale mesocosms, a diverse range of taxa, and different environments. To provide large amounts of data over short periods, NRC recommended prioritizing single-species experiments that represent "best case" scenarios over experiments that examine complex and interactive effects. In addition to experimental work, field-based descriptive data are needed to ground-truth the models and verify the experimental data. The NRC recommended that the descriptive data should focus on sentinel estuaries and that the studies should be conducted repeatedly for a minimum of a 10-year time period. Field surveys aimed at detecting invasions should be conducted in conjunction with ballast water measurements. NRC cautioned that, in the opinion of that committee, while proxy variables can be measured, it is important to not focus on variables that potentially may not represent the risk-release relationship.

The overall goal of this document is to further advance a cohesive binational strategy (United States and Canada) to fill some of these data gaps. This document is an initial step in a progression to develop a cohesive and effective approach, providing an outline and additional detail for various approaches considered in the NRC 2011 document. Provided some or all of the identified data gaps are filled, a second iteration of regulatory efforts to derive ballast water discharge limits designed to protect water quality would benefit from an improved understanding of risk assessment methodologies for evaluating ANS establishment probabilities at various concentrations of living organisms in ballast water discharges (Albert et al. 2013), including at the IMO/USCG standard. Furthermore, these science-based risk analyses could provide support that existing or potential future regulatory limits adequately reduce the risk of ballast-water mediated introductions of invasive species. EPA and the Coast Guard have indicated they would go beyond the IMO/USCG standard if necessary to protect the aquatic environment (77 Federal

Register 17254 2012). Results of risk assessment methodologies will help these agencies better estimate the extent to which existing ballast water regulatory limits are protecting the aquatic environment, or whether, while considering other anthropogenic pathways for ANS, they should consider establishing regulatory limits with more stringent (lower) numeric concentration-based discharge limits.

2. WHICH MODELS ARE AVAILABLE?

Many approaches have been taken to model the risk-release relationship, and it is not possible or desirable here to comprehensively describe all potential models as they have been reviewed in detail elsewhere (Carlton et al. 2011, Lee et al. 2013, Wonham et al. 2013). The intent of this section is to briefly discuss potentially useful models that can use experimental data and characterize population establishment in relation to propagule pressure from ballast water discharges. We focus here on models that have been cited as particularly attractive for understanding the risk-release relationship (Population Viability Analysis and the recently developed Per Capita Invasion Probability models), as well as models that have not been extensively examined elsewhere (Lee et al. 2013). We adopt model categories identified in Wonham et al. (2013).

Generally, attempts to understand the risk-release relationship aim to estimate the functional form and parameters of the following equation (Carlton et al. 2011),

$$P_E = f\,(PP, \varepsilon) \tag{1}$$

where P_E is the probability of a species establishing a self-sustaining population, f is a function, PP is the propagule pressure (number of individuals or frequency of introduction) over a given temporal-spatial scale, and ε is a modifier that accounts for the many variables that can influence the establishment of a species. There is little consensus as to the form of this function for ballast-transported organisms, although by definition it is bound by 0 and 1 and is generally thought to increase with PP.

Similar to any model-fitting routine, available descriptive datasets can be used to estimate the function described by Equation 1. The approach includes first estimating variables (P_E, PP, ε) from available data, then fitting various models to the data and, finally, choosing the best model via a predetermined model selection criteria. Several types of models can be fit to these data, ranging from single-species descriptive models to multi-species mechanistic models, and these are thoroughly detailed in the 2011 NRC report (Carlton et al. 2011). Briefly, the major dichotomy in model types lies between descriptive and mechanistic models. Descriptive models use standard statistical techniques (i.e., linear regression) to characterize the shape of the equation above, and do not attempt to understand *how* the independent variables (PP, ε) influence the dependent variable P_E. Crucially, these models can only be used to predict the relationship within the range of observed values, so they are not useful outside of the limited spatio-temporal scope of the observed data. Mechanistic models, on the other hand, have meaningful parameters that can describe the *process* by which PP and ε alter P_E, and so can be used to extrapolate beyond the observed data.

Mechanistic models will ultimately be the most useful for ballast water management because they enable users to evaluate different management strategies under future scenarios that have not yet been observed (e.g., continued climate change). Fitting these models will sometimes require more data than are available, in which case hybrid mechanistic-descriptive models can be used until more data are collected. Continued progress towards more predictive modeling will require filling critical data gaps. One potentially promising and cost-effective approach to filling such gaps is identifying, collating, and meta-analyzing existing datasets.

2.1 Probabilistic and Dynamic Demographic Models

The following sections discuss various models and their associated strengths and weaknesses.

2.1.1 Population Viability Analysis

Population viability analysis (PVA) uses quantitative methods to predict the likely future status of a population or collection of populations (Boyce 1992, Morris et al. 1999, Beissinger and McCullough 2002). Although originally designed for use in populations of conservation concern and to predict extinction risk, the analyses and fundamental principles derived from these analyses can help us understand the risk of invasive species establishment, assuming there is an inverse relationship between extinction and establishment (Wilson 2000, Drake and Lodge 2004, Andersen 2005). For example, an analysis of how to lower extinction risk to 5% over 50 years is not much different from an analysis of how to increase extinction risk to 95% over the same time period. Similar to risk analysis for threatened and endangered species, PVA-based methods may be used to assess the risk of one particular species establishing itself at one particular site, which may be used to compare the risk of establishment of several potential invaders at one particular site, or to determine the maximum number of arrivals of potential invasive species that is tolerable at a certain risk level (Drake and Lodge 2004, Andersen 2005). When using these methods for invasive species, extinction is the desired outcome rather than something to be avoided.

PVAs range both in method and in the quantity of data needed to parameterize them (Morris et al. 1999, Morris et al. 2002). However, all PVA models consider population growth in a stochastic environment that results in slower growth than might be expected based on mean vital rates (Dennis et al. 1991, Morris et al. 1999, Andersen 2005). PVA approaches can be classified into four general types: count-based, structured, metapopulation, and spatially explicit. Count-based PVAs, the simplest model, require population size data as well as trends in population size over time and assume that vital rates vary stochastically. Structured PVAs use projection matrix models that use demographic data including numbers of individuals in different stages (e.g., age or size categories) in a population. The benefit of this type of PVA is the ability to assess the influence of a particular class on the growth of the population, processes that are potential targets for management (Caswell 2001). For example, projection matrices could determine if it is more effective to manage the different transitional stages of invasions (Morris et al. 1999). Metapopulation PVAs examine the fates of multiple subpopulations to determine whether the rate of establishment of new subpopulations is sufficiently high to counter the extinction of subpopulations. Spatially explicit PVAs simulate the behavior of individual organisms within a given landscape mapped with locations of suitable habitat patches. This model type is the most data-intensive requiring information about birth and death rates of individuals, movement patterns, and degree of isolation and fragmentation of habitat patches.

Most of the information needed to perform most PVAs is count data, in which the number of individuals in a certain population is censused over multiple (not necessarily consecutive) years (Morris et al. 1999). Such data are already abundant for a variety of invasive species and are relatively easy to collect for previously uncensused populations (Brown et al. 2001, Taylor and Hastings 2004). Experimental or count data are used to estimate the

instantaneous growth rate and the instantaneous variance of population growth rate. These two population parameters are then used to simulate the future growth of the population and the population's risk of extinction, including the mean time to extinction and the cumulative probability of extinction at a specific future time (Morris et al. 1999, Beissinger and McCullough 2002). Using experimental data in PVA models allows the dose-response relationship to be quantified under controlled environmental conditions and various environmental scenarios. For example, Kramer and Drake (2010) demonstrated extinction due to predator-driven density-dependent effects on an experimental *Daphnia-Chaoborus* system. Their PVA model revealed that the critical density below which population growth is negative depends on the details of the predator-prey interaction.

In the absence of data, PVA models can also be used to predict a relative change in probability instead of the absolute probability of population establishment (Lee et al. 2013). For example, the Coast Guard predicted the change in the probability of establishment associated with different organism standards (USCG 2012). However, the Coast Guard estimated the population parameter values for μ (mean) and σ^2 (variance) because these values are not widely available in the literature. Experimental data could be used in these PVA models to provide more realistic estimates of these values, allowing the models to predict a more accurate probability of establishment (Lee et al. 2013) and better able to inform management.

As stated earlier, although population viability analysis is used mostly in the recovery of threatened and endangered species, these same methods can be applied in controlling and eradicating invasive species (Bartell and Nair 2004, Neubert and Parker 2004, Andersen 2005). Furthermore, much of the data that are required to perform PVAs for a variety of invasive species already exist. PVA methods can advance the understanding of general invasive ecological principles, help the screening of potential species of concern, and be used for quantitative modeling of risk and for developing invasive species control strategies. Using PVAs in such situations will not only help evaluate the effectiveness of current control strategies and identify areas of improvement, but also help bridge the gap in literature between the role of PVA in threatened species recovery and in invasive species control.

2.1.2 Stock-Recruit Models

The risk-release relationship is also at the core of the applied ecological issues associated with stock-recruitment analysis in fisheries management. Stock-recruit (S-R) models are used to help manage recruitment (the number of offspring that survive) of fish and wildlife populations and are based on the assumption that recruitment is the key factor that most influences the adult abundance (i.e., stock) of many vertebrate species (Hoff 2004a, b, Hoff et al. 2004). This type of model describes density dependence in populations and is used to manage fisheries because it provide estimates of density compensation needed and optimal sustainable yield(Winemiller 2005). The basic S-R model uses only adult stock abundance to predict recruit abundance. When environmental variation strongly influences recruitment, however, S-R models are typically inaccurate; any persistent environmental cause of low recruitment will result in the erroneous appearance that low spawning-stock abundance is the reason (Gilbert 2002). In such circumstances, abiotic factors and environmental variables can be incorporated into S-R models, increasing their predictive capabilities (Crecco et al. 1986, Weber and Brown 2013). The data requirements of S-R models include information about adult population density or abundance,

recruit stage data (juveniles or adult), and sometimes biotic and abiotic variables that affect recruitment.

Few S-R models have been used for invasive species management, although implications from such models could be useful for controlling certain species. For example, Hoff (2004a) developed a S-R model for the Lake Superior rainbow smelt, a nonindigenous species that is currently managed to prevent over-fishing. The results of that study showed that suppressing the number of adults was the most effective way to reduce recruitment, which is of value to managers if they seek to control populations of this species (Hoff 2004a). Determining the stock abundance optimum can be used to control invasive species, assess the feasibility of control efforts, and predict the effects of various levels of control effort. Additionally, simple S-R models that do not account for most of the variability in stock size can be supplemented with abiotic variables. For instance, adding mechanistic variables into a simple S-R model examining the link between invasive insects and climate change improved predictive capacity (Estay et al. 2009). The information provided by these more complex S-R models can not only increase the predictive capabilities for invasive species establishment but also allow the management of other factors such as abiotic conditions.

2.1.3 Epidemiological Models

Because there are many similarities between invasions and epidemics, epidemiological models might also prove useful to invasive ecology. Population growth of invasive species is limited by finite resources such as space and food, and epidemics are similarly bounded by the number of susceptible individuals (hosts). In both cases, limitation increases as a function of the population size (Drake 2005). Additionally, geographic and taxonomic patterns of invasions are strongly influenced by trends in the network that transports them (Perrings et al. 2005, Meyerson and Mooney 2007), just as epidemics are influenced by host movement. Epidemiology can be used to model habitat patches or ports as nodes that are linked by dispersal events. This kind of pathway assessment and network analysis represents a different way of examining the risk associated with biological invasions (Hulme 2009).

Epidemic models are used to describe the transmission of communicable disease through individuals. The most common epidemic models are compartmental, in which individuals in a population are assigned to different subgroups representing specific stages of an epidemic, and transition rates from one subgroup to another are mathematically expressed using differential equations. In the simplest form, Susceptible-Infective (SI) models can have two compartments. This model can then be extended to include recovered (SIR) and exposed (SEIR) subgroups as well as vital dynamics such as birth and death rates.

Epidemiological studies show that the rate of transmission and spread of a disease is influenced by the spatial distribution of the hosts, the susceptibility or immunity to infection of the hosts, and the rates of movement of infected hosts to uninfected populations (Power 1996, Zhang et al. 2000). Similarly to invasions, the likelihood of an invasive species being transported by a particular vessel will depend on the degree of infestation of the hub (how abundant a problem species is and therefore the exposure of the vessel to an invasive species), how susceptible a particular vessel is to being ‘infected’ by an invasive individual, and patterns of vessel movement that influence the connectivity of a given port. Characterizing transport nodes

is important for assessing invasion risks, and the likelihood of a vessel being infected by an ANS depends on the potential source population and the susceptibility of the vessel. Vessels departing a particular port carry with them a community of species that reflect the composition of the species at the mooring port (Floerl and Inglis 2005), and the organisms' size is constrained by the size of the mesh in the grates covering ballast water intake pumps. Busy ports have a high density of susceptible ships, large areas of available habitat for invaders, and high likelihood of receiving infected vessels (Floerl et al. 2009). In addition to evaluating the susceptibility of a particular vessel or type of vessel, knowledge of vessel origin and destinations and information about donor ports and receiving environments is critical to this pathway risk assessment and to determining the rates of movement of infected vessels to uninfected locations. Using risk maps to identify nodes where the initial introduction is most likely to occur could provide guidance for establishing proactive field monitoring methods.

Applying the SIR model to invasive ecology is limited mostly to simulating the spread of marine invaders by hull fouling (Costa et al. 2008, Floerl et al. 2009, Herborg et al. 2009). These models incorporate the variation in a vessel's susceptibility to colonization by the invader and are calibrated based on patterns of vessel movement and maintenance (Floerl et al. 2009). Additionally, the SIR model allows the user to determine the transmission probability (i.e., the probability that a susceptible vessel will become infected at a given location at a given time). After identifying the most susceptible vessels, the model can develop a risk map that identifies areas most at risk for invasions by estimating vessel traffic from geo-referenced monitoring data (Herborg et al. 2009). Because cargo ship traffic is very efficient and ports are closely connected (Kölzsch and Blasius 2011), network analysis of vessel movements could indicate which types of vessels are most likely to introduce invasive species. It would also be possible to model the susceptibility of a recipient region over time. For example, a region that becomes less polluted may become more susceptible to invasions by species previously unable to establish (Carlton 1996). Alternatively, epidemiological methods used to estimate the fraction of the population that has to be vaccinated to halt the spread of an infectious disease (Anderson and May 1991, Hethcote 2000) could be used to evaluate the effectiveness of different management measures by using a treatment technology similarly to vaccination.

2.2 Statistical Models: Per Capita Invasion Probability (PCIP)

One approach to modeling the relationship between ballast discharge and invasion rates is using the per capita invasion probability (PCIP). The PCIP, as explained in Lee et al. (2010), is the "per year probability that an individual non-native propagule discharged from ballast water will become established as a new nonindigenous species in a specified waterbody." Functionally, $PCIP=N_h/(D_h*C_h)$, where N_h is the historical annual number of new invading species per year, D_h is the historical annual volume of foreign ballast water discharge per year, and C_h is the historic concentration of non-native, but not yet established organisms being introduced to a port. Reusser et al. (2013) contains more details of the formation, application, and most critically, the assumptions supporting the use of the PCIP.

The greatest strength of the PCIP approach is that it uses information readily available for many ports and time periods. Because direct measures and estimates of propagule pressure are difficult and costly to obtain, it is assumed the PCIP formulation makes a reasonable proportionate estimate of the propagule pressure, and therefore the risk posed by introduction of

nonindigenous species through ballast. Table 3 of Reusser et al. (2013) describes in detail the potential sources of error and the consequences of errors. Without question, using limited existing data and formulating the PCIP in a way that is connected to our best understanding of the invasion process is helpful and moves science and management forward, resulting in more informed discharge standards. However, three major concerns arise by adopting or focusing management based solely on a simplified approach like PCIP risk assessment:

1. The predictive accuracy and robustness of the PCIP remains unproven.

2. The PCIP does not consider some key factors known to be useful for risk assessment of invasive species (discussed below).

3. Using or adopting the PCIP as a management standard on its own could unintentionally slow implementation of improved risk assessment approaches and provide misleading results.

Predictive accuracy and robustness. As the PCIP is a historical measure, it does not account for changes in the amount and concentration of taxa in ballast water discharged at a port, such as from ships arriving from new origin ports due to changing global trade patterns (Reusser et al. 2013). Studies of the PCIP using established backcasting and forecasting statistical techniques to evaluate predictive accuracy (e.g., see MacIsaac et al. (2004)) or using similar data for other countries (such as Australia) to assess the robustness of the PCIP predictions would greatly bolster confidence in its acceptance as a management-guiding tool. Although measures of relative risk have been used previously to identify shipping ports of most concern of invasion of some species (Herborg et al. 2007), making the additional connection to the invasion rate may be tenuous and is founded on the assumption of a linear relationship of a dose response curve. The robustness of this assumption will have consequences for predictions, such as the claim by Reusser et al. (2013) that "approximately one new species will invade every 10-100 years under the IMO/USCG discharge standard of <10 organisms with a body size of ≥50 μm per m^3 of ballast."

Key factors influencing invasion risk. Habitat matching (Ricciardi 2013) and documented history of invasiveness in other locations (Ricciardi 2003) have been shown to be indicators of a species' potential invasiveness. Neither of these two factors are direct inputs into the PCIP. This would imply that not all connections between two ports are equally risky. Recent studies have revealed that including such factors can result in dramatic shifts in global patterns of predicted invasion risk. For instance, Seebens et al. (2013) demonstrated that ports in the North Sea, despite being highly linked through maritime traffic to other regions—and thus receiving correspondingly high levels of propagule supply—exhibit decreased invasion risk due to low environmental matching with most potential donor regions. When environmental heterogeneity was removed from the risk model, relatively low risk ports such as Antwerp and Hamburg were elevated into the top 20 high-risk category. That study similarly incorporated a measure of "biogeographic dissimilarity" (Tuomisto et al. 2003) to modify the risk model by limiting invasion risk to the proportion of species likely to be alien to a recipient region. The results of that work dramatically illustrate the importance of considering factors other than propagule supply in the risk model. Refining the PCIP to include some understanding of environmental

matching and assessing the proportion of risky nonindigenous species in a sample would likely improve the PCIP (Keller et al. 2011, Rius et al. 2012).

The potential dangers of adopting PCIP. PCIP may represent the current state of understanding and data availability, but adopting this model should also allow for change when improvements can be made. Given that there are key factors known to influence invasion success (see above), and that there is considerable ongoing research into risk assessment and data collection of known factors (i.e., Keller et al. 2011), advancements will undoubtedly be made that could quickly be implemented. Adopting an *active adaptive management plan* would help the process of risk management of ballast-mediated invasions (McCarthy and Possingham 2007) by providing a means to change or possibly replace PCIP as we learn more about best to estimate invasion risk (Grantham et al. 2010).

Even in the absence of perfect or complete information, decisions will still be needed to manage ballast water discharge of ANS. Other sectors of industry and science, such as aerospace, nuclear, and chemical, move forward with research and commerce in the face of significant uncertainties into potential threats to human life and the environment, largely because they have coupled known processes with measures of uncertainty (Bedford and Cooke 2001). Managing invasive species through regulation of ballast water discharge would be well served by accounting for uncertainty, either through traditional experimentation, or perhaps more useful, the incorporation of expert opinion into risk assessment (Cooke 1991). This method has been used for estimating economic damages of invasive species (Rothlisberger et al. 2010), but could potentially, and usefully, influence estimates of PCIP for global shipping pathways or ports—provided that any conclusions drawn are appropriately characterized in terms of the simplified and retrospective nature of its operation.

2.3 Common Data Needs

Advancing our understanding of the risk-release relationship and our predictive capabilities requires using one or more models along with meeting the associated data requirements (Table 2-1). Since many models share some data requirements, the same data can be used in a variety of different models to test for agreement. However, if a preferred model can be identified, data collection can be designed to serve the needs of that model. Since ANS invasions are a result of simultaneous and repeated release of model species, single-population models are unrealistic. However, they may prove useful for modeling particular species thought to be representative, aggressive, or worst-case scenario invaders. Gathering data on target species could provide limits on the establishment probabilities of a much larger set of species for which collecting comprehensive data would be difficult (Wonham et al. 2013). Experimental laboratory studies on target species could be used to determine vital rates required by some models and useful to others. Common model data requirements include instantaneous growth rate and variation along with biotic and abiotic variables that affect population growth. The primary challenge will be to convert these laboratory and mesocosm results into a vessel-scale management threshold because it requires determining how the actual invasion risk from many species in many ships across many locations differs from the predicted risk of one species in a controlled experiment (Wonham et al. 2013).

Table 2-1. Data requirements of different types of risk-release models (categories derived from Wonham et al. (2013)).

Type of model	Data requirements
Descriptive statistical	Estimates of initial abundance (propagule pressure or proxy); estimates of establishment rates (or other establishment-related endpoints); modifying factors contributing to establishment likelihood (e.g., environmental matching)
Biological probability	Estimates of initial abundance (propagule pressure); estimates of establishment frequency; shape parameters accounting for Allee effects; modifiers accounting for effects of abiotic variables
Dynamic demographic	Birth, death and maturation rates; dispersal parameters

3. LEVERAGING EXISTING DATASETS

The following sections discuss existing data on propagule supply and associated potential analyses of these data.

3.1 Overview

A substantial amount of data already exists on propagule supply. A recent literature search conducted for this study identified 66 published references in which animal, phytoplankton, or both types of species were identified in ballast water or ballast sediments. These publications were summarized into a table capturing the following information: brief description of research; ballast water sampling location; origin of ballast water (port, country, or ballast water exchange location); targeted species (virus, bacteria, phytoplankton, protists, zooplankton, invertebrate, fish); number of samples; sampling method; vouchering and preservation methods; time frame of samples; location of samples; public access to data; principle investigator (PI); and PI contact information and affiliation (see Appendix 11.1 for an abridged version of this table). In the 15 cases where papers were unavailable, some information is missing for some of the categories.

Prior to the 1990s, very little research on ballast water was published (see Figure 3-1, Appendix 11.1). Only six papers were found that had sampling periods in the 1973 to 1989 time frame. However, 28 papers were published on studies that collected ballast water or sediment samples from 1990 through 1999, and 32 papers were found that covered samples taken from 2000to the present. Sampling occurred at various locations around the world, although the bulk of studies were conducted in North American ports (37 of 66 references, 56%; Figure 3-2). The sampling design varied substantially, with up to 187 different ballast tanks sampled on as many as 372 ships for each study. A number of studies sampled only one vessel (12 references, 18%) in an effort to determine either the spatial distribution of organisms within a ballast tank or the temporal changes that occur due to the voyage conditions, treatment, or both. The majority of studies sampled more than 10 vessels (36 references, 55%, Figure 3-3). Sampling effort ranged from only a handful of samples to hundreds.

Most publications described sampling methods and how the samples were analyzed or preserved but few commented on whether samples had been vouchered and stored for future use. Sampling methods included net tows (29 references), water samples (20 references), sediment samples (11 references), and wall scrapes (four references), or some combination of sampling methods (15 references). When reported, the majority of samples were preserved in various concentrations of formalin/formaldehyde solutions, Lugol's solution, or a combination of formalin and ethanol. Only nine studies stated that samples had been preserved only in ethanol (concentrations from 70-95%), and only one indicated that samples were stored specifically for future DNA extraction (Doblin et al. 2004). Although publications frequently provided lists of observed taxa, no reference indicated availability of raw data through publicly accessible databases. Based on the available data in the publications, it is unknown if any vouchered specimens exist and, if so, if they are available for future research.

The majority of studies focused on organisms ≥50 µm (52 references, 79%); 27 references identified organisms between 10 and 50 µm and only 15 references identified bacteria

or viruses. Research that was conducted only on ballast sediments concentrated mainly on identifying resting cysts or spores of diatoms or dinoflagellates as well as the resting stages of invertebrates within the sediment (Hallegraeff and Bolch 1991, 1992, Bailey et al. 2003, Bailey et al. 2005). Other researchers simply looked for a single species such as the oriental goby, *Acanthogobius flavimanus* (Middleton 1982), the harmful bloom alga, *Aureococcus anophagefferens* (Doblin et al. 2004), or the bacterium *Vibrio cholera* (McCarthy and Khambaty 1994). Eleven papers examined the efficacy or efficiency of ballast water exchange in preventing nonindigenous species introductions (e.g., Locke et al. 1991, Locke et al. 1993, Drake et al. 2002, Cordell et al. 2009), and one looked at the effectiveness of heat treatment on ballast water (Rigby et al. 1997). Three papers concentrated on the propagule pressure in coastal ecosystems (Verling et al. 2005, Lawrence and Cordell 2010, Briski et al. 2012b). Regardless of the research intent, many publications provide in-depth lists down to the genus and species level of all taxa found within the ballast water, sediment, or simply attached to the inside of the ballast tanks. These data are a resource that can be used to extract the best possible estimates of the variables (P_E, PP, ε) associated with the risk-release function. However, none of the studies provides a direct measurement of ballast PP, which is understandable given the logistical constraints of identifying and counting microscopic organisms in large volumes of water. A more critical gap is that information regarding species richness and abundance is severely limited for these studies due to challenges associated with morphological identification of zooplankton, especially larval forms without recognized species-diagnostic characters. Future sampling efforts may meet these challenges in part by adopting genetic approaches (see Section 8).

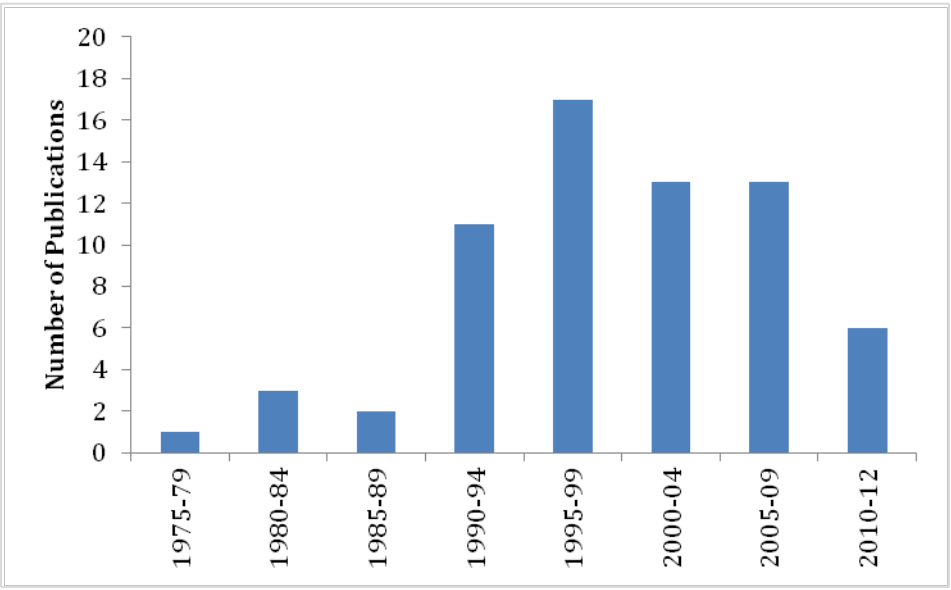

Figure 3-1. Number of publications on ballast water samples per 5-year period between 1975 and 2012

No publications were identified prior to 1975. Note that the final category (2010-2012) is not a full 5-year period.

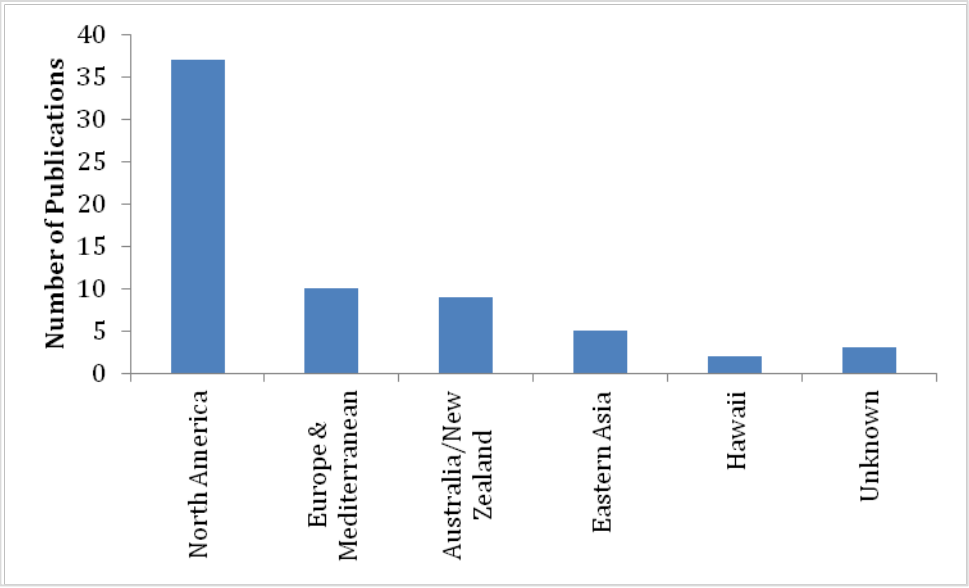

Figure 3-2. Locations of ballast water survey studies for publications between 1975 and 2012

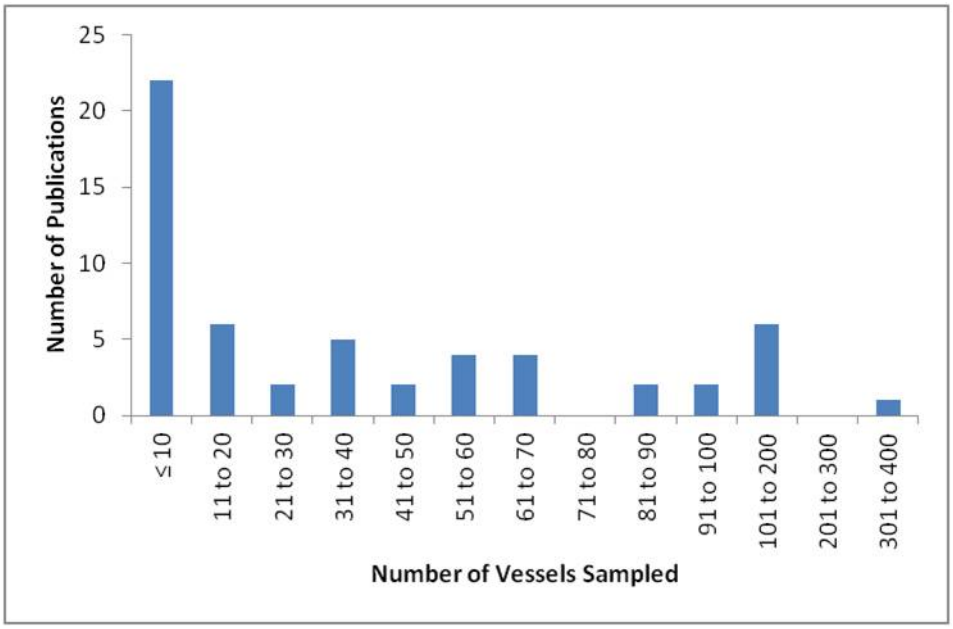

Figure 3-3. The distribution of publications that sampled a given number of vessels

Note that the range of vessels sampled changes at higher numbers (10-vessel bin size up to 100, 100-vessel bin size from 101 to 400). No bar = 0 publications.

3.2 Estimating Propagule Pressure from Descriptive Datasets

PP can be estimated indirectly using arithmetic with a combination of shipping, ballast discharge, and ballast sampling data; see Minton et al. (2005), Verling et al. (2005), McGee et al. (2006), Lawrence and Cordell (2010), Miller et al. (2010) for examples. In general:

$$PP = S * \left(\frac{discharge}{ship}\right) * \left(\frac{N}{discharge}\right), \tag{2}$$

where *S* is the number of ships arriving over a given time period, discharge per ship is the volume of ballast water discharged by ships visiting the port, and *N* is the number of individuals of the species of interest. *S* can be obtained directly from shipping datasets, such as those available from Lloyd's LLC (http://www.lloydslistintelligence.com) or from the National Ballast Information Clearinghouse (NBIC, http://invasions.si.edu/nbic/, USA only). Port traffic patterns vary widely over space and time (Drake and Lodge 2004, Kaluza et al. 2010, Wang and Wang 2011, Ducruet and Notteboom 2012), so this is a simple yet important value to estimate. Continued collection of global ship traffic data is very important for developing predictive risk-release models. Currently, only Lloyd's LLC has comprehensive shipping data at the global scale. A noncommercial, global ship monitoring program could provide a check, backup, and freely available alternative to Lloyd's dataset and would greatly facilitate reaching the goals of this workshop.

The actual volume of ballast discharged per ship ($\frac{discharge}{ship}$) is much more difficult to estimate because it is not typically recorded in shipping datasets nor does it necessarily correlate with shipping traffic volume (Verling et al. 2005, Miller et al. 2010, Wonham et al. 2013). For U.S. ports, the NBIC database can provide ballast discharge data from 1999 to the present. For time periods prior to 1999 or for ports outside the United States, indirect estimates of ballast discharge can be calculated by estimating the size of ships' ballast holds and multiplying by the probability that the ship will discharge a certain percent of its ballast water at a given port. These estimates require knowing the size and types of ships visiting a port (available in the Lloyd's dataset) and using a descriptive relationship between ballast discharge probability and ship type (see Endresen et al. (2004), Verling et al. (2005)). Alternatively, discharge records from ports or ships can provide direct discharge estimates (e.g., Cordell et al. 2009). This latter approach would be considerably more accurate and could be achieved at the global scale as part of the ship monitoring program mentioned above.

The number of individuals per unit volume of ballast discharge ($\frac{N}{discharge}$) will vary by ship, previous ports visited by a ship, and ship voyage conditions. As indicated above, a large number of studies provide direct estimates of taxon density in various ship-port combinations (see above, Appendix 11.1). Some studies have attempted to conduct statistical analyses assessing the importance of various factors in determining propagule density. For instance, Cordell et al. (2009) analyzed ballast from 380 ships arriving in Puget Sound, assessing the contribution of trans-Pacific vs. intracoastal trips, method of ballast water exchange, ship type, and source. They found highly significant differences in propagule supply between foreign (East Asian sources) and domestic (intracoastal) trips, with the latter delivering greater propagule

loads to Puget Sound. However, even within this study, unequal sampling and statistical interdependence of multiple factors reduced power of the analysis and limited inferences regarding sources of variation. Partly as a result of such difficulties, in most studies, even when intensive sampling has been conducted, researchers have not fully characterized sources of variation in propagule supply per unit volume. For instance, Minton et al. (2005) report substantial variation in propagule density among ships carrying foreign ballast into four U.S. ports, but results of analysis to assess factors contributing to this variation have not yet been published. These difficulties are exacerbated when attempting to compare results across multiple studies where different sampling methods have been used (David and Perkovič 2004). Meta-analysis, which is a statistical framework to combine results from disparate studies, would be one approach to best estimate taxon density, its variance, and how it is influenced by sampling method, voyage length, voyage conditions, ballast treatment, and other factors known or suspected to influence organism concentration over the course of a ship's voyage. This descriptive meta-analysis could provide initial estimates of $\frac{N}{discharge}$ under a variety of scenarios and could also help prioritize future descriptive and experimental studies. However, the potential utility of even a meta-analytic approach to leveraging existing data will be constrained by limits on the accuracy and precision of species richness and abundance estimates (i.e., the problem of morphological identifications), as well as the fact that existing data may be of diminishing relevance in a policy era in which historical patterns of ballast water delivery are disrupted by exchange, treatment, and other management actions.

Ultimately, the logistical constraints of ballast sampling indicate that density of most species from most ports will remain unknown via direct estimates. Alternative indirect estimates of $\frac{N}{discharge}$ must consider biogeography (is this species present in the ports visited by the ship?), abundance in port waters (how abundant is this species and how does abundance vary by season?), swimming behavior (can this species actively avoid entrainment in ballast?), ship ballast dynamics (did the ship uptake ballast at this port?), and the effects of ship and voyage conditions on density over time. A species' presence or abundance in a particular port could be estimated from one of the several online species distribution databases (i.e., Global Biodiversity Facility www.gbif.org, Ocean Biogeographic Information System www.iobis.org) or from natural history knowledge, assuming that such information is limited in part by undersampling of some regions, an lack of taxonomic expertise to identify some organisms, and the dynamic nature of species distributions. Swimming speed (and therefore likelihood of avoiding entrainment) could be inferred from body size or morphology (Bellwood and Fisher 2001). Ballast uptake probability could be estimated indirectly via any of the three ways mentioned for ballast discharge. Finally, estimates of the effects of voyage length or conditions for similar taxa could be applied (see MacIsaac et al. (2002) for an example). Altogether, this method could provide a crude estimate of species density for any particular taxa on any particular voyage. Estimates from this indirect method could be tested against current and future empirical data to refine the method and estimate error rates.

Finally, when estimating *PP* for a given spatio-temporal scale, the structure and dynamics of the shipping network should be considered. Over a given time scale, the total *PP* discharged to a port will equal the sum of *PP* from each individual ship that discharges ballast to the port during that time. The total *PP* of a ship will include some combination of ballast uptake from

each port visited on its voyage plus effects from the intervening voyages between ports. Such "carry-over" or "stepping-stone" effects need to be considered to accurately estimate *PP* and can be achieved by modeling spread through the shipping network (see Floerl et al. (2009), Kaluza et al. (2010), Keller et al. (2011), Kölzsch and Blasius (2011), Seebens et al. (2013 for examples)). For instance, Seebens et al. (2013) recently applied risk models that incorporated details of the global shipping network, and demonstrated that explicitly considering invasion risk associated with ports prior to a ship's last port of call dramatically altered the overall global risk profile. Further developing these risk models to capture "stepping-stone" effects and incorporating more accurate propagule pressure estimates will provide useful tools for policy analysis.

3.3 Estimating Modifiers (ε) from Descriptive Datasets

Several factors unrelated to *PP* can modify the probability of a particular species establishing in a particular port. As reviewed in the 2011 NRC report, these factors include, but are not limited to, species traits (life history, invasion history, genetic variation, niche breadth, dispersal, mobility, environmental match to introduced location) and the physical and biotic environment of the new port (habitat landscape, hydrodynamics, disturbance regime, interactions with resident biota).

Environmental match and prior history of successful invasion have been found to have the largest and best predictors of establishment probability in other systems (Hayes and Barry 2008, Bomford et al. 2009). Estimating the effects of these two factors for ballast-mediated invasions should therefore be of highest priority. For both effects, knowledge of successful and failed introduction events, along with the associated *PP*, would provide the most straightforward estimates. Environmental match could be calculated by using widely available climate data (i.e., Worldclim (www.worldclim.org), the World Ocean Atlas (http://www.nodc.noaa.gov), or the Global Ports Database in Keller et al. (2011)), and prior invasion success could be determined by searching online invasive species databases (i.e., The United Nation's Fisheries and Aquaculture Department Database on Introductions of Aquatic Species (DIAS; http://www.fao.org/fishery/dias/en) or Smithsonian Environmental Research Center's National Exotic Marine and Estuarine Species Information System (NEMESIS; http://invasions.si.edu/nemesis/)). For example, Seebens et al. (2013) assessed environmental matching using temperature estimates drawn from the World Ocean Atlas and salinity estimates based on data from Lloyd's Register Fairplay (http://www.port-guide.com); they also considered "biogeographic dissimilarity" (Tuomisto et al. 2003) as a factor influencing risk of invasion. That study found that including these factors resulted in dramatic shifts in global patterns of predicted invasion risk. This dramatic illustration of the importance of considering factors other than *PP* underscores the inherent complexity of accurate risk models.

The most straightforward statistical methods to estimate any modifier's effect on ANS establishment require data on successful and *failed* introductions across different levels of the modifier and *PP*. Unfortunately, failed introduction events are rarely, if ever, recorded, and associated *PP* for any event can be difficult to estimate (see above section). It is therefore unlikely that we will be able to obtain quantitative estimates for the effects of these or any of the other modifiers mentioned above. It may yet be possible, however, to assume that high environmental match and prior invasion success increase P_E, and then use this information to

create relative risk rankings for different species, ports, and routes (e.g., see Herborg et al. (2007)).

3.4 Estimating Establishment Probability (P_E) from Descriptive Datasets

Establishment probability, P_E, is the chance that a species will establish at a new location to which it has been introduced. It can be directly calculated by tracking the outcomes of new introductions but, to our knowledge, there are no such descriptive studies available specifically for ballast organisms. Alternatively, P_E can be indirectly estimated from descriptive datasets that track both new establishments and *PP* over time. There are several datasets of new invasion records over long time periods: San Francisco Bay (Cohen and Carlton 1998), Coos Bay (Ruiz et al. 2000), Puget Sound (Ruiz et al. 2000), the Laurentian Great Lakes (Mills et al. 1993, Ricciardi and MacIsaac 2000, Ricciardi 2006), Europe (Gollasch 2006), Chesapeake Bay (Ruiz et al. 2000), the Mediterranean Sea (Galil 2009, Zenetos et al. 2012), and Port Phillip Bay, Australia (Hewitt et al. 2004). However, these represent only a small subset of the world, and, notably, they often cannot distinguish between multiple potential introduction vectors. Finally, and most problematic, these datasets lack concurrent *PP* estimates, so they are not useful for mechanistic models.

Given the lack of shipping data from which to derive *PP* estimates for these historical datasets, the most plausible approach for quickly filling the P_E data gap is to intensively sample for both new invaders and *PP* over shorter periods of time. A few studies using traditional taxonomic methods (Lawrence and Cordell 2010) appear to have the necessary data to do this. However, the logistical constraints of sampling and identifying highly variable plankton communities prevent wide application of this approach. Recent developments in environmental DNA methods could potentially make plankton sampling and new invader identification easier and more standardized, which would facilitate a coordinated effort to monitor for new invaders in many ports at once (see Section 6). These data, together with a concurrent global ship and ballast-monitoring program, would greatly facilitate developing predictive, mechanistic risk-release models that could be useful for ballast water management.

3.5 Recommendations for Analysis of Existing Datasets

Existing data are, unfortunately, inadequate to inform the choice of models best suited to understand the ballast water risk-release relationship (Carlton et al. 2011). To our knowledge, no formal meta-analysis has been conducted on available ballast water sampling data. Such analysis could be conducted rapidly (within the first two years of the proposed research) and could yield insights into the factors driving variation in propagule pressure that could be used to improve existing models, to identify appropriate sampling strategies for collecting additional data, and to inform selection of future models. We have identified a large number of studies that describe propagule supply in various contexts (Appendix 11.1); a first step toward analyzing these data would be developing a coordinated database including raw data (organism counts, sampling strategies, ballast source and destination, etc.) from these and other existing studies. Such a database could be adopted as a central repository for similar data collected in the future, and would thus provide a framework for standardizing reporting and an opportunity to leverage more effectively efforts conducted independently throughout the world. Limitations on the value of historical data analysis for future assessment of the risk-release relationship (e.g., limits on

estimates of species richness and abundance, changing policy climates resulting in novel patterns of ballast water delivery) suggest that the greatest value in developing and maintaining such a database will likely be in the collation and standardization of future sampling data. This database could have enormous value in terms of identifying temporal patterns of propagule supply related to shifting regulatory climates (see section 9.2 on ancillary benefits). More directly, it would provide a single, easily accessible dataset for future meta-analysis. It is important that the structure of such a database be defined collaboratively and publicly, to increase the likelihood that future data collection efforts, even if conducted outside the auspices of the coordinated research effort described here, can be seamlessly incorporated into statistical analytical frameworks.

4. EXPERIMENTAL STUDIES

The following sections discuss the purpose of experimental studies and existing experimental studies, as well as associated issues and design recommendations.

4.1 Aims of Experimental Studies

The study of biological invasions seeks to explain and predict the likelihood of establishment with obvious implications for management agencies and regulatory policy. However, predicting invasions is one of the most complicated problems faced by ecologists because the invasion process consists of a series of stages, each with its own suite of complexities and uncertainties (Ruiz et al. 2000, Kolar and Lodge 2001). Additionally, questions about whether regulations are sufficiently protective and future development of standards may need to rely on quantitative models that require parameterization and validation. Not only can we advance our understanding of invasion ecology through experimental studies, but quantitative predictive models can also be parameterized by experimental results.

Invasion biology presumes that there is a quantifiable relationship between the number of individuals of a given species released into an environment and the probability of its eventual establishment (Ruiz and Carlton 2003). Although biologists assume that the risk-release relationship is positive, the precise nature of the response could vary greatly over species and environments (Ruiz and Carlton 2003, Lee et al. 2013). Invasion biology is based on a synthesis of occurrence records, and the discovery rate of invasive species is not unbiased or a reliable surrogate for the rate of invasion (Ruiz and Reid 2007). Due to this deficiency in theoretical knowledge of the risk-release relationship, management agencies typically base regulations on expert opinion and subsequently infer a risk-reduction level. Ecological principles and information about specific species traits can be used to suggest conditions under which the probability of invasion would most likely be maximized (Kolar and Lodge 2002, Carlton et al. 2011), but the uncertainty associated with a given release rate can differ by 100-fold (Lee et al. 2010). Experimental studies could provide data to help resolve this high complexity and uncertainty in the risk-release relationship. Therefore, one goal of experimental studies is to inform the theoretical understanding of invasion dynamics that is essential to developing appropriate discharge standards.

Experimental studies are also needed to parameterize and validate risk-release models. In addition to having a theoretical understanding of the risk-release relationship, models require reliable measurements of the variables used for predictions. Additionally, determining the uncertainty associated with model predictions requires validation of models that currently has not been done due to a lack of data and bias in sampling protocols due to a lack of standardized methods for the data that are available. Experiments are valuable in their ability to control propagule pressure and expose target communities to many different conditions while measuring the outcome. Experiments are a cost- and time-efficient way of obtaining these critical data. Experiments using high-impact or commonly released species could be used to parameterize models, and those models could then be used for informing a conservative discharge standard (Carlton et al. 2011). Fitting risk-release curves to experimental data could allow management agencies to quantitatively predict the effect of certain discharge standards on invasion risk.

Regulations on the concentration of viable organisms in ballast discharge water are established around different organism size classes (IMO 2004a, USCG 2012, Albert et al. 2013). We focus in this report on organisms greater than or equal to 10 µm in size. These include two IMO/USCG size classes: organisms ≥50 µm, which include holoplanktonic organisms, adult and larval fishes, and larval states of benthic organisms; and organisms measuring ≥10 µm and <50 µm, which include most phytoplankton. The IMO/USCG standard for the ≥50 µm size class represents a considerable reduction in the concentration of organisms, decreasing from an estimated several thousand to less than 10 organisms per m^3 of ballast water (Minton et al. 2005, Lee et al. 2013). The focus on these size classes is due to the larger availability of theoretical, empirical, and experimental data (Lee et al. 2013), the lack of baseline information available for microorganisms (Ruiz et al. 2000), and the belief that holoplankton are less likely to be polyvectic (Carlton et al. 2011). This narrow focus reduces the sampling effort, increases the probability of detection by concentrating on species that can be accurately identified and counted in ballast samples, and provides the most likely clear signal for analysis of the risk-release relationship.

4.2 Past and Ongoing Experimental Studies

Although limited experimental colonization studies have been conducted in aquatic systems, it is nevertheless clear that such studies have advantages and disadvantages (Chadwell and Engelhardt 2008, Bailey et al. 2009, Carlton et al. 2011). Controlled experimental studies of the risk-release relationship in ballast discharge are particularly difficult because either the volume used is too small or the density of aquatic organisms is too high to simulate realistically the dynamics of ballast release. Previous work can inform experimental design efforts to focus on data gaps and standardized methods that allow for consistent comparison across research studies.

There are three experimental scales for assessing invasion risk as a function of propagule pressure: benchtop, mesocosm, and field studies. It should be noted that the scale of an experiment depends on biological factors relevant to the systems and phenomena under consideration. Benchtop-scale experiments are defined as those that could conceivably be conducted in most laboratory settings where water tanks could be placed and manipulated. Mesocosm-scale studies could conceivably span a much wider range of volumes, but typically require dedicated space and plumbing. Field studies are conducted within ambient environmental conditions and are conducted outside laboratory facilities with limited or no ability to contain water.

Benchtop-scale experiments are considered useful in the absence of larger-scale enclosures or appropriate field sites and when space or time is limited (Carlton et al. 2011). The water volumes required in benchtop-scale experiments may also be easier to sterilize after experiments are completed than those in a mesocosm. Effluents from benchtop experiments with invasive species propagules can, for example, be sterilized with chlorine bleach and then dechlorinated if necessary before being released into aquatic environments or wastewater systems.

Although field studies offer the most realistic conditions for any ecological study, they present challenges to researchers in terms of controlling environmental factors. In field studies of

invasive species, the risks of introducing experimental species to uninvaded waters must be managed in some way. Some studies have eliminated this risk by choosing field sites where focal species have already invaded (e.g., Porter-Whitaker et al. 2012, Cockrell and Sorte 2013).

Mesocosm-scale studies offer many benefits as they are more easily controlled than field studies, and if planned carefully, are more realistic than many benchtop-scale experiments. Mesocosms have been used to study community effects of focal invasive species and the effects of climate on native and invasive species populations (e.g., Cataldo et al. 2012, Fey and Cottingham 2012, Reynolds and Bruno 2012, Cockrell and Sorte 2013), and this experimental scale shows much promise for establishing quantitative relationships between propagule pressure and invasions. The volumes of water needed for mesocosm studies are potentially highly variable and will require treatment if effluents could contain invasive propagules. Although the need for dedicated space for large mesocosms and relevant numbers of replicates pose challenges, mesocosm-scale studies show the most promise for determining propagule pressure/invasion risk relationships. A number of facilities currently exist that can conduct such experimental work, and new systems continue to be added in North America and elsewhere. For example:

1. As of 2013, the Great Ship's Initiative (GSI) is researching how to generate empirical information on the risk-release relationship. GSI's goal is to help determine credible methods for assessing the risk of establishment and providing scientifically sound experimental data across a range of taxa relevant to the Great Lakes ecosystem.

2. The University of Rhode Island administers two facilities capable of invasive ecology mesocosm studies. The Island Marine Ecosystem Research Laboratory (MERL) in Narragansett, Rhode Island, houses 14 tanks for use in experimental studies that can contain up to 13 cubic meters of water each (one of these tanks is temporarily unusable due to deteriorated support structures (Maranda 2013)). At this time, effluents from this facility are not treated, but effluents may not require treatment if the target species under investigation are native or naturalized in the local environments. The other facility, the Luther Blount Aquaculture Laboratory, houses a quarantine-capable pathology laboratory comprising five wet labs, a live feed room, and phytoplankton culturing facilities (Baker 2013). This laboratory space can accommodate benchtop-scale studies of various sizes and also includes larger tanks that range between 189 and 946 liters. Effluents can be passed through ozone treatment before disposal if necessary.

3. The United States Geological Survey (USGS) operates two facilities in the Puget Sound region of Washington State at which invasive species propagules could be studied in mesocosms under the Western Fisheries Research Center (WFRC). WFRC comprises five research facilities, and the facilities at Seattle and Marrowstone could support invasive propagule research. Marine organisms could be studied at the WFRC Marrowstone facility in tank sizes of 60 liters or more. Effluents from this facility are treated with chlorine before entering a freshwater discharge pond. The WFRC facility in Seattle can support freshwater research as well as research on pathogens, and also treats effluents with chlorine before release to the Seattle wastewater system (Smith 2013). Both facilities support

current research programs and expertise in aquatic invasive species (WFRC 2013).

4. Bodega Marine Laboratory in Bodega Bay, California, is a field station of the University of California at Davis. This field station has received funding from the National Science Foundation to install an ultraviolet (UV) water treatment facility that would allow mesocosm-scale studies on marine invasive species (Cherr 2013). Prior to receiving this funding, Bodega Marine Laboratory provided chlorination treatments for effluents of studies of shellfish pathogens and invasive species. The new facility is intended to expand and provide infrastructure for research on invasive species and their propagules and to replace the current chlorine treatment system and dechlorination procedures that pose potential risks to local environments (Cherr 2013).

The important features needed for a facility to support the suggested experimental research include space for multiple mesocosm tanks, water treatment capabilities, and expertise in aquatic invasive species identification. The mesocosm tanks used in invasive species ecology have ranged from 60 L to more than 10,000 L. The size of the tank depends largely on the species of interest and the treatment factors under consideration (see Section 4.4). Experimental facilities should also have the ability to quarantine the system or treat the effluent to prevent accidental release of nuisance species or they should use only local biota.

4.3 Issues of Scale and Extrapolation

Experimental systems and mesocosms can be useful as representations of the complex ecological interactions in nature, exposing key processes through simplification. However, to develop predictive models needed for regulatory decision making, we must learn how to interrelate phenomena acting on different scales (Levin 1992). The most commonsensical solution would be to conduct studies at the whole-ecosystem scale, but that approach to studying invasive species can be infeasible. Therefore, benchtop-scale experiments (microcosms) can be an inexpensive way to build knowledge about more complex systems and can even reveal insight into the effect of scale if conducted across a gradient of scales (Carpenter et al. 2010).

Although there are many definitions of scale in science, we refer to scale as the spatial and temporal dimensions of a pattern or process of interest (Cumming et al. 2006). For example, data can be collected at a spatial scale of a whole country or at the smaller scale of an individual site. Additionally, data can be taken at large temporal scales over long-term time periods or shorter time periods. Some patterns may not be evident at certain scales since different processes are likely more important at different scales, and multiple mechanisms may be needed to explain patterns on all scales (Levin 1992, Schneider 2001). For the risk-release relationship, there is typically both a spatial and temporal mismatch between propagule pressure and establishment data (Bradie et al. 2013). Spatial scale mismatch results from using propagule pressure proxies that focus on high-level measurements of trade (i.e., shipping activity, Ricciardi (2006)) whereas establishment typically refers at a more site-specific scale. Temporal scale mismatch is common because identifying rare invasions requires extended time periods to see establishment trends (Simberloff 2009), whereas propagule pressure data are typically more short-term and contemporary (i.e., one year) (Bradie et al. 2013).

For experimental studies, the two ends of the scale spectrum include small local experiments and whole ecosystem (i.e., whole lake) experiments or surveys. Whole-ecosystem experiments are hard to replicate, expensive, difficult to execute, and it is not often possible to control important variables, which confounds the interpretation of the results (Schindler 1998). For invasion ecology, whole-ecosystem experiments are further complicated by the fact that intentional introductions of non-native species would be unreasonable, unethical, or not practicable. Whole-ecosystem studies are typically too crude to provide detailed causal understandings of mechanisms.

On the other hand, mesocosms provide the fine-scale resolution and control needed to isolate certain conditions and understand specific mechanisms (Schmitz 2003). The challenge is to translate experimental or mesocosm results, which represent simplified conditions, to patterns in the ecological systems of interest to management and policy. Experiments can, among other things, reveal complex interactions, identify fundamental mechanisms, and allow for the parameterization of models (Carpenter and Kitchell 1988, Pace 2001). However, experiments are typically done at scales far below the scale of interest requiring extrapolation or estimations of a value of a variable outside its measured range. For example, most common in invasive ecology are single-species experiments and models (Leung et al. 2012). These small-scale mesocosm experiments can be used to forecast and fine tune whole-lake experiments or conversely, experimental results can be verified using subsequent whole lake surveys (Schindler 1998). Extrapolating from small scales to broad scales may allow for greater generalizations but eliminates finer details that are irrelevant for producing the observed patterns at the larger scale (Levin 1992). Direct extrapolation from mesocosms to whole ecosystems is often questionable and potentially erroneous (Schindler 1998, Pace 2001). For example, phytoplankton responses (23 of 60 species) in enclosure experiments were quite different to those in whole-lake experiments, with enclosure experiments only correctly predicting whole-lake responses for 34% of the taxa tested (Carpenter and Kitchell 1988). Even results from small whole-lake experiments needed correction to be extrapolated to larger lakes (Fee and Hecky 1992, Fee et al. 1996). Conversely, a bacterial-primary production relationship derived from mesocosm tanks was the same as the relationship derived from field studies in marine and freshwater systems (Hobbie and Cole 1984, Cole et al. 1988).

Although scale has been widely considered in ecological process examinations (Levin 1992), it has been less so in invasive biology. Researchers have compared experimental data and field surveys in invasive ecology to evaluate the impact of invasive species (i.e., Matsuzaki et al. 2009) and to evaluate the detection of invasive species (i.e., Britton et al. 2011). Few studies have extrapolated results of mesocosm studies of the risk-release relationship to field data, and the complexity of accurately measuring propagule pressure for ballast introductions makes integrating field and laboratory data even more difficult. One example is Gertzen et al. (2011) who experimentally controlled propagule pressure using different stocking rates of invading spiny water flea in mesocosms and then integrated these results with field data to generate a population model that estimates the probability of establishment. Additionally, Bradie et al. (2013) developed a risk-release curve for aquarium fish species imported for trade and then determined if this model could explain the historical survey data of species that have successfully established. Despite the ability to accurately make quantitative predictions about establishment in the absence of species-level trait information, valid prediction could only be expected when the propagule pressure data for a given introduction pathway are relatively consistent across time

(Bradie et al. 2013). While little work has been done to ground-truth invasive species models, lessons learned from extrapolating causes and consequences of other ecological phenomena can easily be applied.

The accuracy of predictive models, and their value in regulatory contexts, depends on whether scale issues are handled properly. The value of individual-based models relies on a quantitative understanding of how individuals respond to a range of environmental conditions combined with a quantitative description of the environmental conditions in the area of interest (Huston 2002). Successful extrapolation of mesocosm experiments to ecosystem-level theory requires: (1) time and spatial scales large enough that the observed response is not transient, (2) realistic simulation of the natural system, and (3) the noise associated with the variation in time and space is small compared to the response (Pace 2001). Theoretical studies are needed to suggest mechanisms and explore relationships and environmental surveys to quantify patterns at broad scales (Levin 1992). Understanding the risk-release relationship will require experimental work done simultaneously with field surveys and synthesizing information from multiple scales. The more consistent results are across all approaches, the more confidence there should be in the inferences made from such results (Carpenter 1998). It is not a matter of choosing the correct scale but rather recognizing that change is going to take place between different scales (Levin 1992, Schneider 2001). This point has further implications for needed synthesis to reach management-level scales, which are generally even larger than ecosystem scales, as further discussed in Section 9.

4.4 Design Recommendations

Modeling the relationship between the number of organisms in ballast water and the likelihood of invasion requires an understanding of the density-dependent effects and the stochastic aspects of population growth, among other things. The risk-release relationship is most likely specific to a given species (or type of organism) and/or a specific environment. Therefore, experimental studies need to be tailored to specific taxa and a consideration of environmental, hydrodynamic, and biotic factors applied during the design phase. The scale of the potential experiment might drive some of these decisions, and the scale of an experiment will also rely on the treatment factors under consideration.

Because of the regulatory need for a very broad understanding of the nuances of interactions among propagule pressure and environmental factors, it is necessary to ask how establishment probability varies across environmental conditions, information that can also be obtained from experiment studies (Wonham et al. 2013). The experimental design procedures should consider both abiotic and biotic factors. Favorable abiotic conditions with ample food will allow for a worst-case scenario analysis (i.e. highest probability of establishment). However, realistic consideration of the variation at the discharge site is important as well. For example, how does the water chemistry vary over the time scale of the experiment or over the time period required for sampling?

As recommended by Carlton et al. (2011), experimental studies should focus on taxa that represent worst-case scenarios for invasion (see Section 7.1). This approach will enable parameterization of models aimed at identifying the most precautionary limits on the risk-release relationship, which will ultimately be useful for setting boundaries for risk management

decisions regarding discharge standards. Experimental work should also focus on single species. Although invasion is a complex process and very rarely, if ever, involves introduction of single species, the experimental approach is able to gather substantial data under controlled circumstances, allowing robust statistical analysis of factors impacting establishment success. Given the likely need to conduct such studies at the mesocosm level and the limited availability and relatively high cost of such experiments, resources are best applied to understand the factors underlying the risk-release relationship for single species. Once experimental approaches have been refined sufficiently to confidently parameterize models for single worst-case target species, and only then, should efforts addressing multiple species invasions be considered.

Experimental systems must be able to support establishment of target species at some relevant inoculum level, within an experimentally tractable time frame. These constraints are but nontrivial. Many of the most attractive target taxa may not already have protocols for rearing in the laboratory, or may not be culturable in the laboratory, for that matter. That means that establishing positive experimental controls (conditions under which establishment is known to occur) is critical. Having no such controls will make it impossible to derive robust inferences regarding the relationship between establishment rates and inoculum density. Given rearing requirements and life cycles of most organisms in the ≥50 µm size class, mesocosm experiments capable of running time courses of weeks to months will be absolutely necessary to assess establishment risk at management-relevant levels of propagule supply. Fortunately, facilities capable of supporting such experiments already exist (see Section 4.2).

Data collection should focus on temporal sampling of population density as well as experimental endpoints (e.g., success and failure of establishment) so that experiments can support development of both statistical models of establishment success *and* dynamic demographic population models (Wonham et al. 2013). The latter requires time course information on population density. This type of data is also valuable because it will improve our understanding of how best to assess population abundance of targets (e.g., sampling strategy, live counting, genetic methods). In this context, "establishment" must be carefully defined, ideally in such a way that is meaningful to management and can be translated to port-scale surveillance efforts. At the very least, establishment should require completion *in situ* of one full reproductive cycle.

Whenever possible, experiments should be conducted at multiple scales (e.g., micro and mesocosm). Working effectively at multiple scales will likely entail experimental work with phytoplankton, which typically fall in the ≥10 µm to <50 µm size class and which could be particularly valuable for assessing scale effects. Phytoplankton could easily be scaled from benchtop microcosms (e.g., glass beakers) to benchtop midicosms (e.g., 10-20 gallon aquaria) to full mesocosms (hundreds to thousands of liters). This would allow investigators to rapidly explore a wide range of environmental conditions on the smaller scales, and then verify the results at more realistic larger scales, as well as determine scaling functions that would assist in extrapolating these findings to real world scenarios. As noted above, pursuing megacosm experiments—conducted in enclosed or sheltered natural systems—has many challenges. While initial efforts at experimentation should focus on benchtop and mesocosm scales, it is worth considering future megacosm scale experiments, particularly for their potential value in validating parameterizations based on smaller scale studies. Unfortunately, studies at this scale entail additional challenges, such as the potential for damage to experimental systems through

natural event or vandalism (Bailey et al. 2009), limits on the ability to introduce certain organisms (or, alternatively, the difficulty of recreating preinvasion conditions in already invaded systems), and limits on the ability to replicate across multiple trials.

Ideally, if multiple facilities could coordinate their experimental studies, it would increase confidence in the reliability of experimentally derived parameters and limit concerns regarding differences in equipment differences and quality control across different sites. This would require multiple facilities to not only replicate experiments, but also adopt standardized experimental protocols designed to minimize potential biases between sites and to ensure that variation in results can be attributed to relevant biological differences.

5. DEVELOPING A SHIPBOARD SURVEILLANCE STRATEGY

The following sections discus important aspects of developing a shipboard surveillance strategy, including sampling intensity, vessel and pathway classification, sampling approaches and cost.

5.1 How Much Sampling is Needed?

Ballast water regulations center around defining acceptable limits of living organisms, yet the relation between the number of establishments and the numbers of introduced organisms are unknown. Moreover, even propagule pressure and colonization pressure are likewise difficult to estimate, given the data available. This problem is made even more challenging because the receiving ecosystem is complex, there are multiple sources of variability, and there are only limited resources available. In reality, it is not feasible to measure all sources and densities of propagules, or even identify all species being introduced. However, we assert that with the appropriate models, we can make valid statistical inferences about the numbers being introduced, even with limited sampling. While it is premature to quantify the exact amount of sampling required, we can at least identify the major considerations. Here, we focus on estimating propagule and colonization pressure, rather than on their linkage with the probability of establishment. Below, we discuss three considerations: 1) sources of variation and uncertainty, 2) issues of scale and optimal allocation of sampling effort, and 3) statistical approximation of propagule and colonization pressure.

First, because resources will be limited, it is important to identify the sources of variation and uncertainty so we can identify areas to allocate different amounts of effort and how data collected can potentially be used to infer introductions. These sources of variation include both *intra-ship* and *inter-ship.* Intra-ship variation occurs because organisms will likely not be uniformly or randomly distributed within the ship. The number of species and their abundance detected in any given ship will depend upon the extent of sampling (i.e., for both the species richness and the numbers of organisms, intra-ship uncertainty reduces asymptotically to zero as sampling increases).

Inter-ship variation may occur in several ways and for several reasons. First, ships may differ in the types of organisms introduced (species identity), as well as the species richness and the number of species. These differences may be due simply to stochastic processes, but also to factors such as source location, ship (vector) type, treatment technology, time period (season and year), and distance travelled (e.g., Sylvester et al. 2011). These may be used as predictor variables to more finely resolve the variables of interest - colonization pressure and propagule pressure. Because not all ships or ports will have been measured, there will also be inter-ship uncertainty as well. Ports may differ substantially in the types of traffic visiting them, although much of this variation may potentially be explained by source, vector type, and distance travelled. The receiving environment (at ports) will also be important to link this information to the probability of establishment (Leung et al. 2012).

Second, because resources will likely be limited, we should consider the optimal mix of sampling at different scales: 1) intra-ship, 2) inter-ship within ports, 3) between different types of ships or from different source locations, and 4) between ships in different ports (or different

geographical locations). The amount of effort required to add an additional sampling unit to each of the four scales identified above would determine the optimal mix (e.g., multiple samples within a ship may be less expensive than multiple samples across different ships), the amount of variation at each scale, and the reduction of uncertainty afforded by each additional sampling unit (see below). Optimal solutions may be obtained via algorithms such as dynamic programming.

Third, as mentioned above, we are unlikely to sample exhaustively or find all species being introduced. However, while determining the identity of every species introduced may be unfeasible, it may be possible to statistically estimate the propagule pressure and colonization pressure more generically. These general estimates can then be input into a vector-level species establishment model (e.g., Bradie et al. 2013). Potentially, the model can use rarefaction curves to extrapolate species richness and use total counts of living organisms and dominance-diversity curves to determine the distribution of propagule pressures across species. Note that unanswered questions remain about how best to integrate rarefaction curves, total propagule numbers, and dominance diversity curves across the different scales, plus their associated uncertainties. From a theoretical perspective, computer simulations could be used to evaluate the consequences of different sampling approaches (i.e., different protocols for distribution of sampling effort at multiple scales) under various scenarios, to define which inferences are valid and to identify potential generalities.

5.2 Classification of Vessels and Pathways to Inform Choice of Surveillance Targets

The various pathways by which an organism may enter a body of water, including vessel type and mechanism of introduction (e.g., hull fouling or ballast water), are important factors in designing a long-term surveillance program for identifying nonnative species. A variety of vessel types, including international merchant vessels, coastal domestic merchant vessels, and domestic lake vessels (Lakers), all pose an invasion risk to ports in the United States; however, the risk is not solely related to vessel type alone, but also to vessel origin, frequency and volume of ballast water discharge, vessel route, and transit success survivorship (Verling et al. 2005). Various studies have examined the effect of many of these variables on invasion risk (Verling et al. 2005, Ruiz and Reid 2007, Cordell et al. 2009, Simkanin et al. 2009, Bailey et al. 2012, DiBacco et al. 2012).

In the past, international merchant vessels posed a high level of risk of introducing ANS. Ballast water from international ports is now required to be exchanged or flushed in the open ocean, which considerably reduces potential propagule supply (Cordell et al. 2009, Bailey et al. 2012, although see Drake et al. 2002). Due to the factors discussed below, coastal domestic merchant vessels may now pose a greater risk for species introduction than international merchant vessels. Regarding hull fouling, Bailey et al. (2012) suggests that these vessels are a primary concern to the Great Lakes St. Lawrence Region (GLSLR) because short domestic voyages are expected to transfer more healthy organisms than would long foreign voyages. The risk from ballast water from these vessels is also greater than from international merchant vessels for a number of reasons:

- First, there is an inverse relationship between duration of voyage and propagule survival (Lavoie et al. (1999); Verling et al. (2005) and Simkanin et al. (2009)).

The vessels are more likely to spread native ANS or established nonindigenous species (NIS) than to introduce new NIS from foreign sources (Carlton and Hodder 1995, Lavoie et al. 1999).

- Second, ballast water exchange is not required for coastal domestic merchant vessels by U.S. federal regulations and is not regularly practiced, sometimes due to insufficient time in transit between ports. As suggested by Ruiz and Reid (2007), this presents a loophole in the framework established to protect against the transfer and spread of NIS in U.S. waters. Two studies demonstrate the increased risk that domestic merchant vessels impose on U.S. coastal ports. Cordell et al. (2009) found that densities of high risk taxa were consistently and significantly higher from U.S. domestic trips dominated by tank ships carrying ballast water from California, and lower in samples from trans-Pacific trips dominated by container ships and bulk carriers with ballast from Asia. These results were more than likely a result of the dense and diverse NIS assemblages present in California and other U.S. west coast estuaries and the comparatively short transit times. Similarly, Simkanin et al. (2009) found that the overlap of ANS among port systems on the west coast varied between 3% and 80%, with the largest overlap occurring between San Francisco Bay and Los Angeles/Long Beach. The results from this study suggest that intracoastal ballast water should be further examined as an invasion pathway, especially because short-sea shipping is heavily promoted today.

- Lakers, or domestic vessels that operate exclusively in the GLSLR, are responsible for 95% of ballast water discharges in that region and appear to be the most important transport pathway of hull fouling and ballast-mediated NIS in the region (Bailey et al. 2012). Regarding hull fouling, these vessels travel shorter distances within routes of similar latitudes, further increasing propagule introduction. As mentioned above, propagule survival is inversely related to voyage duration, suggesting these vessels may have a higher risk of propagule supply than international vessels (Lavoie et al. 1999, Simkanin et al. 2009).

As a result, it has been suggested that future efforts to reduce dispersal of NIS in the GLSLR focus on domestic vessels, particularly Lakers (Bailey et al. 2012). Though Lakers will not introduce species from foreign sources, they are likely to spread native nuisance species and or established NIS once these organisms are introduced to the GLSLR (Carlton and Hodder 1995, Lavoie et al. 1999). By developing categories of vessels that register important factors determining variation in propagule supply, a vessel classification scheme can facilitate design of ballast water surveillance programs that allow robust statistical inferences of overall patterns of propagule supply to a recipient system. Surveillance efforts targeting Great Lakes ports, for example, should consider the potential importance of Laker traffic, while marine ports should explore variation in propagule supply between overseas and domestic sources (e.g., Lawrence and Cordell 2010).

5.3 Comparison of Ballast Sampling Approaches

There are two current approaches for collecting ballast water samples: collection from the ballast tank (in-tank) or from the piping system used to transfer or discharge ballast water (in-line; IMO 2008). Ballast tanks can be very large (thousands of cubic meters), multilevel tanks often containing ladders and other internal structures, and their design configuration varies from ship to ship. In-tank sampling approaches rely on accessing ballast tanks through hatches, manholes, sounding pipes, or venting pipes. Through these access points, ballast water is pumped out into nets or filters, or equipment such as nets or water column sampling devices are lowered into tanks to collect samples of ballast water organisms.

The advantage of using in-tank sampling techniques is that it allows researchers to perform studies of biota over the course of a voyage using multiple sampling methods (e.g., conducting vertical tows of plankton nets through hatches or using pumps to collect water samples through sounding tubes). In addition, organisms are subjected to very little physical manipulation that could cause mortality during sampling (Cangelosi et al. 2003), and by accessing the ballast tank directly, the sampling team can collect sediment samples from the tank bottom (Dodgshun 2003). Limitations to in-tank sampling techniques include limited access to tanks, tank conditions affecting the ability to collect samples, and the amount of time necessary to collect samples. Access to ballast tanks may be unpredictable, unsafe, and crew-time intensive (Cangelosi et al. 2003). Tank conditions, such as low volumes of water, tanks filled to capacity, or obstacles within tanks, can affect the ability to collect samples (Hamer 2003). Most importantly, the samples of organisms collected will not be representative of the organisms present in the tank (or volume of interest).

In-tank sampling approaches assess potential introduction of organisms rather than point-of-discharge conditions (David and Perkovič 2003). In-ballast sampling techniques are the most historically common sampling method. However, because organisms are stratified within ballast tanks (e.g., Murphy et al. 2002, First et al. 2013), samples collected using in-ballast methods will not be representative of the water of the volume of interest (here, the tank). Furthermore, to evaluate treated ballast water—which contains sparse numbers of organisms—using the Poisson distribution is warranted (Lemieux et al. 2008, Miller et al. 2010). If samples are not collected in a manner that is representative of the volume of interest, using the Poisson distribution is invalid.

In-line sampling techniques involve installing equipment to collect samples directly from ballast water pipelines (e.g., Richard et al. 2008). Access points include the main ballast pipeline (before or after the pump, or at the discharge point) or firefighting system pipelines (Cangelosi et al. 2003, David and Perkovič 2003). The USCG Final Rule requires sampling ports to be located "as close as practicable to the ballast water management system overboard outlet prior to the discharge point to determine concentrations of living organisms prior to discharge." (USCG 2012) Prior to sampling, a piping alteration is made to allow for sample collection directly from the ballast pipeline. Equipment such as pitot tubes, catchment tubs, filters, and nets are used to collect ballast water organisms (e.g., First et al. 2013).

Strengths of in-line sampling approaches include the ability to characterize conditions at discharge, evaluate treatment system performance, and replicate sampling across ships. Notably, , if sampling is done correctly (using the proper sample port diameter, which is sized for the

diameter of the ballast main and flow rates, and the water is well mixed), the sample collected will be representative of the volume of interest (e.g., a tank or a portion of a tank). Taking samples from the discharge pipeline under conditions of fully developed turbulent flow means organisms cannot avoid sampling equipment (Cangelosi et al. 2003) and the sample represents organism condition, concentration, and composition upon discharge to the receiving system (Cangelosi et al. 2003, Dodgshun 2003). In-line equipment can be installed both before and after treatment, to evaluate the performance of a ballast water treatment system, or on several ships to allow for comparisons across vessels (Cangelosi et al. 2003). The IMO G2 guidelines recommend using the in-line sampling approach to characterize effectiveness of ballast water treatment (IMO 2008). In addition, research suggests that in-line sampling approaches are the only means to collect samples appropriately (Richard et al. 2008, US EPA 2010), and are the least expensive and most consistent methods (Cangelosi et al. 2009). Limitations of in-line sampling approaches include having to install sample ports (by cutting and welding of the ships' ballast pipes) in advance, the lack of information on spatial distribution of biota within the tank (though smaller organisms' distribution may be determined by collecting whole water samples at different depths), and need for equipment that must be designed and positioned to minimize bias and hydrodynamic effects within pipes. In-line sampling must take place during intake or discharge of ballast water and is often carried out in the engine room, which typically requires crew assistance (although in-tank sampling through a hatch also requires crew assistance) (Cangelosi et al. 2003, Dodgshun 2003). Collection piping must be designed to minimize bias in capture of entrained particles (Richard et al. 2008) and regardless, outcomes may be altered by hydrodynamic effects within the pipes (Cangelosi et al. 2003, Cangelosi et al. 2009). In-line sampling may result in a potentially distorted sample collection due to equipment that causes mortality or selects organisms of certain shape or size; however, Cangelosi et al. (2009) found no difference in organism survival, densities, or diversity between the most effective in-line sampling method and most effective in-tank sampling method. Likewise, in samples collected using an in-line sampling device (a 'filter skid', with housings containing filter bags) or a plankton net as a ballast tank was discharged, First et al. (2012) did not find significant differences between living organism concentrations using the two means of sample collection, and the community structures were similar.

In-line sampling approaches are historically less prevalent than in-tank sampling; however, in-line approaches are becoming more popular due to innovations in method design, the recognition that representative samples are required to provide statistically sound estimates of living organisms, and due to the changing regulatory climate, with an emphasis on shipboard BWMS validation testing and compliance monitoring. For example, the Coast Guard ballast water discharge regulatory standards mandate that an access point for in-line sampling be available for vessels required to meet the numeric discharge limits (USCG 2012). In-line sampling at the point of discharge represents the point of potential ANS introduction itself, and is suitable for conducting a risk assessment of biological invasions and determining compliance to regulatory standards (David and Perkovič 2003).

5.4 Recommendations for Shipboard Surveillance

Analyzing existing ballast sampling data (see Section 3) would refine understanding of the variance in propagule supply that has already been sampled. Published analyses have demonstrated that the data across multiple factors—ship type, trip duration, source region,

seasonality, etc.—are highly variable, showing that there is no strong correlation between propagule supply and currently available measures of vessel traffic, including ballast water volume (Wonham et al. 2013). Unfortunately, it is impracticable—and likely will remain so into the foreseeable future—to determine directly the propagule supply associated with any meaningful subset of the ballast water entering a system. Even determining the propagule supply associated with a single vessel pathway (movement between a single source and single recipient port) would be beyond the capacity of any research program, due to the sampling effort required. The primary aim of shipboard surveillance as described here is to sample enough ships and account for enough proportion of the variance across multiple factors known to influence propagule supply to characterize the probability distribution of propagule densities per vessel and the overall delivery to the entire ecosystem. Analyzing existing ballast datasets should be sufficient to reasonably assess the level of effort required to obtain such statistically robust estimates. Knowing the number and types of vessels entering a system and the types of major source ports, and accounting for seasonal variation, it should be possible to determine the number of ships needed to develop an overall estimate of propagule supply within a predefined confidence interval. For large target ports with complex patterns of vessel traffic, this sampling effort would likely exceed 100 ships per year.

This long-term research effort will require intensive sampling of three to six target ports (see Section 7.2 for section criteria), ideally the same systems as those selected for surveillance of recipient environments. While it would be desirable to sample a larger number of ports to capture variation due to biotic and abiotic characteristics of recipient systems, funding limitations might favor more thorough sampling at fewer sites over less intensive sampling at more sites. First, strategically selecting target systems might enable much of the variation in propagule supply to be observed at the regional scale (e.g., variation in source population density, seasonal variation, length of voyage), a snapshot of which can likely be taken at a single recipient port. Second, it will almost certainly be easier to standardize within sites than between them, and limiting sampling efforts at fewer sites can thus reduce managerial overhead. Finally, intensive sampling efforts at single sites allow the best opportunities for a relatively complete temporal assessment of propagule supply to those ports, thus increasing the usefulness of the research effort in terms of testing the impacts of policy changes (i.e., the shift to numerical discharge standards). In this context, it is critically important to develop a comprehensive picture of propagule supply patterns to at least one of the selected target ports during implementation of the recently established discharge standards.

In-line sampling approaches can be readily standardized and are likely to be implemented more broadly in the future; therefore, they should be adopted as the standard approach for sampling. These approaches also have the benefit of assessing propagule supply in a context as close as possible to that through which it is delivered to recipient systems. The drawback of these approaches is that they will increase the expense of the overall effort and require rapid initial investment to retrofit vessels, or identify those that can already be sampled. Once these steps are taken, however, future sampling should be relatively easy to standardize across ships both within and between study ports. Cangelosi et al. (2001) provide detailed recommendations for implementing in-line sampling, including guidance for cost-effectively utilizing financial and human resources.

Although the NRC has recommended stratified random sampling design for ship surveys (Carlton et al. 2011), we currently do not know how to best sample ships to cost-effectively estimate the overall propagule supply into a system (see Section 5.1). As mentioned above, an approach to assess the optimal stratification of sampling within ships, between ships within ports, and between ships in different ports, should be adopted, taking into account the estimated increased accuracy of more accurate propagule supply estimates using more sampling units and the expected costs to add them. Analyzing existing data (see Section 3.2) could add insight by providing additional information on the importance of different factors in determining variation in propagule supply, along with estimates of the level of sampling effort required to adequately account for that variation. Getting access to samples able to support that variation (e.g., Lawrence and Cordell 2010) will also be challenging. Currently, it is not clear how many sampling events might be needed to reliably estimate overall propagule supply. However, hundreds of such events would likely be required, with repeated sampling over multiple years spanning important policy changes (e.g., the shift to numerical discharge standards). Fortunately, this scale of effort is not unprecedented. For instance, Cordell et al. collected over 800 ballast samples between 2001 and 2012, at least half of which have been analyzed ; see also Cordell et al. (2009), Darling (2013c) and Lawrence and Cordell (2010).

It is also important to recognize the need to initiate this effort rapidly. The United States is beginning a second large-scale experiment in ballast water management, this one associated with the move to numerical discharge standards. Such a move, assuming discharge standards are reliably met, will reduce the inoculum density in vessel discharges and should reduce invasion rates attributable to ballast water discharges. Quantifying that reduction, however, will be challenging, particularly in light of the inability to quantify invasion risk from ballast water discharges today.

5.5 Cost Considerations

We are aware of no published record providing total costs for any large-scale ballast sampling effort. In the case of surveillance programs based on in-line sampling, significant costs such as investment in ship modification and reusable operational and biological sampling equipment, must be borne up front. Cangelosi et al. (2001) estimate that one-time costs include $1,500 for initial ship inspections, $2,000-$5,000 per ship for installation of sample ports, up to $2,000 for biological sampling equipment, and $45,000 in reusable operational equipment. Overall one-time costs will depend on the number of ships being sampled and the number of expected sampling events. These estimates do not consider costs associated with staff time or staff travel to sampling sites, which can be considerable.

Costs associated with analyzing biological samples will depend on the type of analysis required. Between 2001 and 2012, researchers analyzed 658 ballast samples of a total of 870 collected during that period from Puget Sound, at a total cost of approximately $235,000 (Darling 2013c). This works out to approximately $357 per sample for standard morphological analysis. Adding molecular genetic analysis would clearly incur additional costs (see Section 8.4). Estimates for the Puget Sound work probably reflect some of the efficiencies obtained by coordinating taxonomy teams to analyze large sample sizes using standardized approaches. However, it is anticipated that representative samples collected via in-line sampling may present additional challenges compared to the net sampling adopted in the case of Puget

Sound. In particular, samples may contain more organisms, and samples may thus require additional analysis. It is reasonable to estimate that standard analysis of biological samples for a 10-year sampling program taking 100 samples per year (total of 1,000 samples) would cost approximately $500,000, and that overall costs associated with an intensive ballast water surveillance effort could be as much as $1 million per target port.

6. DESIGNING A LONG-TERM PORT SURVEILLANCE PROGRAM

The following sections discuss important aspects of designing a long-term port surveillance program, including general considerations, sampling and analytical approaches, and cost.

6.1 General Considerations

Designing a coordinated recipient system surveillance program aimed at informing the risk-release relationship must satisfy a number of criteria: it must effectively detect rare taxa; it must generate data that are comparable across surveys within the same system and across multiple target systems; it must be statistically robust, such that species' presence or absence and the likely number of established ANS can be estimated; and it must be as cost-effective as possible. Substantial effort has already been applied to gathering records on the presence and establishment history of non-native species that have been detected in North American coastal systems (Mills et al. 1993, Carlton and Hodder 1995, Mills and Sommer 1995, Ruiz et al. 2000, Holeck et al. 2004, Wonham and Carlton 2005, Fofonoff et al. 2009). A similar approach has been used in many other regions to compile information on non-native species. Unfortunately, with this approach, most records come from synthesizing available literature, instead of from standardized measures. As a result, the information has been assembled from diverse sources (including research, citizen reporting, etc.) with little or no coordination or standardization among studies in terms of targeted taxa, sampling strategies, or analytical approaches. Although this provides some useful information on non-native species, it is not particularly useful for surveillance or to evaluate risk-release relationships, because (a) these efforts have failed collectively to satisfy the above criteria and (b) there is no clear way to correct for the spatial or temporal variation in data quality (Ruiz et al. 2000).

To control for bias in historical data, standardized surveys have been undertaken in various coastal regions. Most of these have been applied initially to examine spatial patterns of invasion, but also to evaluate temporal patterns. For example, Australia and New Zealand have both implemented standardized national efforts to survey ports for the presence of invasive marine species of concern (Hewitt and Martin 2001a, Inglis et al. 2006a, Sliwa et al. 2009a). In the United States, continental-scale surveys of standardized surveys have also been implemented to detect non-native species and compare spatial patterns of non-native species richness for sessile invertebrate communities (Ruiz and Hewitt 2002).

Such prior efforts can be adopted as models for designing appropriate surveillance approaches to further understand the ballast water risk-release relationship. It is clear that the methods and approaches needed for surveying non-native coastal biodiversity is sufficiently understood, across diverse habitats types and taxonomic groups. The main challenge is in the specific design for surveillance and the model(s) used to characterize the relationship with ballast water discharge (i.e., propagule supply).

We recommend aggressively pursuing design of surveillance strategies for three to six target ports that satisfy the selection criteria identified in Section 7.2. This should include San Francisco Bay, Chesapeake Bay, and Tampa Bay to capture different biogeographic regions and vessel origins and to build on significant, ongoing baseline surveys (of plankton and benthic

communities) being conducted for non-native species by the Smithsonian Institution at each location; including Duluth/Superior Harbor would further expand biogeographic coverage and build on growing surveillance efforts in that system. As with ship surveillance, we favor greater effort in fewer target systems over more diffuse effort over a large number of targets. Concentrating effort at fewer sites will increase cost efficiency, and comprehensive understanding of the relationship between propagule supply and establishment at even one site will vastly improve our ability to assess various models of the risk-release relationship. Furthermore, high-intensity surveillance efforts at few target ports will better allow collection of data that inform future surveillance efforts, for this or other purposes (see ancillary benefits, below). Agencies interested in ANS monitoring typically face a trade-off between broad surveillance of numerous potential sites of ANS incursion and sufficiently high-intensity sampling at single sites to allow robust statistical inferences. For the research effort under consideration, we feel that precedence should be placed on the latter of these two considerations.

At each of the chosen target ports, we recommend pursuing a two-pronged approach to surveillance: targeted assessment of presence or absence and distribution of a small set of taxa (see Section 7.1) and non-targeted overall assessments of biodiversity. While most established surveillance programs have singled out target species based primarily on risk assessment (Hewitt and Martin 2001b, Inglis et al. 2006b, Campbell et al. 2007, Sliwa et al. 2009b), this will be only one of the considerations in choosing appropriate target species. Ideally, the research effort described here would be able to comprehensively assess non-native species establishment in the target system. This means that detection efforts must range far beyond detecting pre-selected targets. Results of other surveillance efforts suggest that standardized sampling approaches are capable of such “non-target” detections (Inglis et al. 2006a). However, it remains unclear the degree to which such approaches can provide comprehensive accounts of non-native biodiversity. In particular, it may be difficult to identify *a priori* target species that will prove most informative in understanding the risk-release relationship. Most of the investment in long-term surveillance should be aimed at looking at total diversity and focusing on sampling key habitats and key taxonomic groups that use ballast water as a vector. We recommend assessing total diversity in samples by both applying traditional approaches to taxonomic identification and incorporating novel methods that should vastly increase the depth of taxonomic resolution and allow simultaneous, cost-effective processing of increasingly higher numbers of individuals (see Section 8). In addition, we recommend adopting statistical methods to estimate the richness of non-native species, even when not detected. If appropriate sampling design can make inferences of species diversity sufficiently statistically robust, total numbers of established non-native species may possibly be estimated, even if their identities are unknown.

6.2 Design of Sampling and Analytical Approaches

6.2.1 Examples of National-Scale Standardized Surveillance Efforts

Both Australia and New Zealand have implemented standardized port surveys for marine invasive species as components of their biosecurity programs. In addition, a parallel program has been established in the United States for over 10 years, although it is only recently being used by state and federal agencies to assess marine biosecurity programs. These surveillance efforts probably represent the most mature and intensive such efforts conducted on a national scale

anywhere in the world, and they can provide useful guidance for designing effective surveillance efforts into the future.

The Commonwealth Scientific and Industrial Research Organization (CSIRO), Australia's national science agency, published a technical report describing general protocols that agencies involved in the Australian National Ports Survey should use (Hewitt and Martin 2001b). The goal of this program is to establish an informational base of the current distribution and abundance of introduced species in Australian harbors and coastal waters. The report reviews and updates the general protocols developed by the Center for Research on Introduced Marine Species (CRIMP) in 1996. It updates the survey standards accepted and ratified by the Australian Ballast Water Management Advisory Committee (ABWMAC) and the Research Advisory Group (RAG) for purposes of baseline introduced species surveys of Australian ports. The report describes survey design guidelines and standardized methods with a targeted approach to limit costs and increase consistency in the face of challenges due to the number of survey sites (72 first ports of call) and number of participating surveying agencies. Additionally, the report provides guidelines on sample handling and archiving, maintenance and availability of archived samples, reporting, and considerations for future monitoring and resurveying.

The Australian National Ports Survey aims to determine the distribution and relative abundance of introduced species and provide baseline assessments of introduced, cryptogenic (of unknown origin), and native species. However, due to the geographic scope of the survey and taxonomic diversity, the program recognized two important challenges: limited funds and consistency between many unique locations and multiple participating agencies. Strategies to address these challenges are incorporated into the standardized protocols.

To increase cost efficiency, the survey design targets pest species and at-risk habitats, uses sampling techniques that minimize sample volume, and leverages local knowledge. Because sample analysis is expected to be the largest survey cost, the protocols focus on target species to minimize this cost. Targeted species include those in one or more of the following categories:

1. The species is listed on ABWMAC’s schedule of international pests.

2. The species is recognized as a major pest in overseas ports that may be expected to colonize Australian ports.

3. The species is known to be present in Australian ports but not yet with pest status.

This pre-selection exercise aims to reduce hours spent on taxonomic analysis for identification and informs spatial planning of the field survey efforts.

Field survey efforts are focused spatially to maximize cost efficiency. Sampling methods must ensure comprehensive coverage of habitats and provide both presence/absence information and semi-quantitative indices of abundance. Sampling is concentrated near introduction areas and habitats most likely to become colonized by the target species. This maximizes the probability of capturing rare, unestablished populations, which are assumed most likely to be found near the point of inoculation.

Selecting appropriate sample techniques and leveraging local knowledge saves additional costs. Sampling techniques that do not produce large quantities of samples or require long sorting periods are selected whenever possible. Exceptions are made when there is a high probability that introduced species are present in a specific habitat, the species are cryptic or not readily recognized, or there is no other sampling technique that will effectively sample that habitat. In addition, leveraging local knowledge increases design efficiency. Each survey begins with a public awareness campaign that identifies recent changes that indicate new species have arrived and includes guided field efforts as to where to sample and what techniques to use based on an estimated timeline of colonization.

Baseline surveys provide a snapshot of the current distribution and abundance of introduced species in a specific port. The duration for which survey results are representative depends on the specific port. The following criteria are recommended in identifying an appropriate resurveying period: 1) the risk of missing the arrival and establishment of a new pest species balancing the upfront survey costs, 2) the known background rate of invasion, and 3) the known speed of specific species colonization. At a minimum, the report recommends resurveying every 3-5 years. Surveys must be conducted often enough to identify invasive species early enough to provide options for eradication.

The New Zealand Biosecurity Council also published a technical report describing the design and trial of a targeted surveillance program of marine pests (Inglis et al. 2006a). This surveillance program was designed to monitor specific target species at high-risk ports in New Zealand. Seven species were identified for surveillance based on their presence on the New Zealand Register of Unwanted Organisms, their known significant impact on ecosystem or economic value, their ability to survive in conditions prevalent in New Zealand coastal waters, and their current lack of establishment in New Zealand waters (with the exception of Asian kelp (Lodge et al. 2012)). The study focused on eight harbors that had been previously identified as high risk due to the volume and pattern of international shipping within them, the availability of suitable habitat for target species, and their history of invasion by nonindigenous species (Inglis 2001). Strategies and methods for cost-efficient detection were developed and evaluated, including using habitat suitability index (HSI) models. Lessons learned from this survey planning and implementation may be applicable to survey efforts in the United States.

The choice of field survey methods was driven by the assumed need to sample a large number of locations in each harbor to increase the likelihood of detecting founder populations of introduced species. Survey methods were selected based on their effectiveness at capturing the target species when present, cost and ease of sampling, level of impact on native marine environments and species, and the safety of field personnel, the general public, and property. Conventional ecological survey techniques were found to be too labor-intensive, nonspecific, and focused on enumeration rather than detection. Prioritizing rapid detection increased the number of locations that could be sampled on each survey. Selected methods included epibenthic sled tows, seastar traps, box crab traps, crab condos, shoreline searches, diver searches, and drop cameras.

The targeted surveillance program was operated on a fixed budget (see cost considerations, below). The budget provided enough for a team of six to survey each harbor four times, for up to six days each, using six methods, on two vessels. Sampling episodes were

conducted six months apart to capture seasonal conditions. With such budget and time constraints, the program sought to guide the field surveys using preexisting spatial datasets on habitats and key environmental variables in local ports and harbors. However, due to limited availability of such datasets, the survey designers supplemented with hydrodynamic and HSI modeling. Based on the known distribution of habitat for the target species and the model output, each harbor was divided into three to four strata (e.g., head/entrance) that reflected broad environmental gradients and the concentrations of particles (i.e., organisms) simulated in the model. Approximately 60% of the sampled locations were chosen from strata where moderate to high weighted mean concentrations of organisms were predicted.

The first field survey gathered information to determine targets for the number of sites sampled with each technique in each harbor. An average time taken to obtain samples using each method was established. Suitable habitats within each harbor were identified and sampling effort distributed between measurement techniques accordingly. This information was used to adjust expectations for number of sites sampled with each technique in each harbor.

Several of the lessons learned described in the report may be used to improve future survey efforts. First, limitations of the HSI model use strategy were recognized (Inglis et al. 2006b). While the HSI models had an overall high predictive success when used to narrow the search area while increasing the probability of encountering an incursion, application was not possible where scarcity of data prevented robust model development. Furthermore, the HSI model failed to detect at least one significant new invasion. Three main factors were theorized to explain this failure: inaccurate portrayal of artificial structure habitat, surveying during dormant periods of the species' lifecycle, and surveying at inappropriate depths to maximize encounters with other species in the same locations.

Additionally, the report made several recommendations regarding model calibration data, species-specific techniques, and the feasibility of survey design for containment rather than eradication. Because the utility of HSI models depends on the amount and quality of information used for calibration, the report recommends collaborating with regional authorities to collect quality environmental data for each harbor. This includes collecting a broad-scale survey of salinity during the survey itself, spatially interpolating sediment and habitat cover observations to fill in areas where such data are not otherwise available, and including multiple nodes of introduction for larval dispersal based on actual shipping activities. The report also advocates developing species-specific survey techniques, including molecular probes and chemical attractants.

Inglis et al. (2006a) specifically challenge managers to define the purpose of active surveillance programs, to determine the amount of survey efforts needed, and identify the appropriate incursion responses. This note is important in the current context, as an effort aimed primarily at early detection of high-risk target species will likely incorporate different design elements than one aimed primarily at providing a more comprehensive picture of the overall non-native biodiversity established in a recipient system. Indeed, the Australia and New Zealand surveillance programs have been somewhat limited in their ability to provide comprehensive overviews of coastal bioinvasion patterns (Inglis et al. 2006a), and it is likely that efforts aimed more broadly at estimating establishment rates will require additional investments.

In the United States, the Smithsonian Institution has designed and implemented standardized surveys to establish baseline data to (1) compare spatial patterns of native versus non-native species richness across bays and ports in the continental United States, and (2) test for temporal changes in new invasions over time in response to changes in vector activity (including management actions) and environmental changes. A primary goal was to use methods that could be replicated (in space and time) to allow comparisons that controlled for habitat type, history, season, and other factors that can influence community composition. This program was launched in 1999 and has surveyed over 30 bays in the continental United States as well as additional bays in Central America and a few other global regions.

The initial focus of these surveys was hard substrate habitats in high salinity waters of bays and estuaries, using a stratified, random sampling design (Ruiz and Hewitt 2002). More specifically, these focused on sessile invertebrate communities, using settling panels as passive collectors that are colonized and analyzed for species occurrence. Use of settling panels controls for age and substrate differences that affect community composition and also allows the salinity, depth, and distribution of sample locations to be standardized. To date, approximately 10,000 panels have been retrieved throughout the country, resulting in over 100,000 voucher specimens that have been identified to characterize community composition and non-native species occurrence on these panels.

At several sites, these surveys are now being replicated over time to evaluate temporal change in non-native species richness and community composition. In addition, the scope of surveys has been expanded at three bays to include macrozooplankton communities and genetic analyses. For the latter, Smithsonian has adopted traditional bar-coding of vouchers to confirm species identification and build a bar-code library as well as metagenomic approaches to assess the entire community composition (see Section 9 for further discussion of these methods). The Coast Guard and the state of California are supporting the recent expansion of these surveys as a first step to evaluate the performance of biosecurity measures that aim to reduce invasions associated with ships (both ballast water and hull biofouling).

6.2.2 Recommendations for Long-Term Port Surveillance

The above examples demonstrate that substantial guidance already exists on designing effective surveillance efforts aimed at determining non-native diversity and detecting novel invasions. These programs have been developed based on well-established principles of coastal ANS surveillance developed to optimize cost-effectiveness and detection probability for rare taxa. Over the past decade, substantial progress has been made toward such optimization. For instance, Campbell et al. (2007) review relative efficacy of five different types of survey approaches implemented in over 100 surveillance efforts worldwide. Additional studies have further assessed efficacy of different sampling methods and strategies (Trebitz et al. 2010, Hoffman et al. 2011), and have recommended approaches that enhance sensitivity of detection and allow statistically robust estimation of “encounter functions” relating sampling effort to the likelihood of detection for rare non-native populations (Hayes et al. 2005).

For surveillance efforts in the current research context, we recommend designs based on a combination of the CRIMP protocols (Hewitt and Martin 2001b) and passive sampling methods (e.g., settling plate deployment), along with genetic analysis to assess community

composition. The CRIMP protocols were designed to maximize the likelihood of detecting introduced species, and have proven effective at detecting even small populations of introduced species within large sampling areas (Campbell et al. 2007). They accomplish this in part by expending most effort where introduced species are most likely to be, for instance, by using tools such as habitat suitability indices and particle dispersion modeling to design stratified sampling schemes (Inglis et al. 2006a, Inglis et al. 2006b). Such approaches take advantage of the aggregated character of many early-stage coastal invasions (Inglis et al. 2004), and they have been demonstrated to reduce the effort required for detecting non-native populations (Hoffman et al. 2011). In addition, the CRIMP protocols adopt quantitative methods that facilitate comparisons among sampling sites and between surveys, which is a critical consideration for a coordinated effort to assess ANS establishment rates (Campbell et al. 2007). CRIMP surveillance also reduces the requirement for taxonomic expertise in the field, transferring it instead to post-sampling analytical teams (as opposed, for example, to rapid assessment surveys, which rely heavily on field taxonomy). Distributing the taxonomic burden is appropriate for a coordinated surveillance network that may be able to leverage centralized taxonomic analysis across multiple field collections.

Applying the CRIMP protocols to surveys for assessing overall establishment patterns should factor in modifications to potentially improve the capacity to comprehensively assess non-native diversity. For instance, many of the invasive taxa identified during Australian port surveys were “rediscoveries” of introduced species already documented in the literature (Sliwa et al. 2009b). While accounting for such species is critically important for a complete picture of establishment history, the cost benefit of such rediscovery must be taken into consideration. In addition, sampling design intended to maximize cost efficiency of detecting target species often depends on assumptions about target distribution that may be incorrect (Hayes et al. 2005). Thus, focusing on sampling areas with expected high density of non-native species (based on habitat suitability, proximity to ballast discharge, or other metrics) runs the risk of failing to detect species that violate our assumptions about likely habitat selection or dispersal from initial points of introduction.

Due in part to these considerations, we recommend supplementing the CRIMP protocols with passive sampling devices. Passive sampling has been used frequently to measure the distribution of non-native species in highly invaded ports, often with a focus on high-risk fouling taxa (Taberlet et al. 2012, Tang et al. 2012, Vettraino et al. 2012, Wood et al. 2013). By choosing appropriate artificial substrates and adopting appropriate statistical models for tempo-spatial patterns of deployment, passive sampling can be a cost-effective method of targeting various coastal communities in a way that is not destructive (allowing a wider variety of post-field analytical approaches) and integrates establishment patterns over a known time scale (Campbell et al. 2007). These tools are a relatively cost-effective means to supplement CRIMP surveys, and may enhance surveillance in areas under-sampled by the CRIMP approach (i.e., areas in sampling strata representing lower *a priori* likelihood of high-density non-native taxa).

Generally, we advocate an explicitly adaptive approach to surveillance design. This approach will be facilitated by tight coordination within and among surveillance efforts and established protocols for formal project review and oversight throughout the lifetime of the program (see Section 9). Detection likelihoods should be frequently assessed and encounter functions should be estimated using rarefaction (species accumulation curves) and other

established methods (Hayes et al. 2005, Inglis et al. 2006a). Sampling strategies should be refined as necessary to account for new information on recipient system environmental or biotic parameters. The adaptive design approach will be particularly important for incorporating new technologies (e.g., genetic methods) that have yet to reach full maturity in the context of early detection and monitoring.

6.3 Cost Considerations

Surveillance for rare nuisance species is expensive. This is even true when considering surveillance for a single target species of concern; for instance, one study estimated that to achieve 95% probability of detecting infestation of an introduced viral pest in New Zealand plum orchards would cost over $2 million (Ganev and Braithwaite 2003). Therefore, effective national-scale surveillance for multiple invasive species requires substantial investment. Inglis et al. (2006a) estimate that implementing a surveillance program across eight New Zealand ports, aimed at providing 90% detection probability of early incursions of seven target ANS, would cost nearly $12 million. Longitudinal studies are, of course, even more expensive. A proposal for a 10-year marine invasive species monitoring effort in Puget Sound, Washington, estimated overall cost at approximately $5.4 million. Table 6-1 provides several additional estimates of cost for intensive, large-scale surveillance efforts.

Campbell et al. (2007) provide the most detailed estimates available for costs associated with various marine surveillance programs. While costs depend on the survey type, estimates across all types ranged between approximately $3,000 and $4,200 per sampling site, indicating that average costs are likely not widely different among different sampling approaches. In some cases, these estimates already account for efficiencies gained by standardization and coordination across target sites. Such efficiencies depend in large part on whether the bulk of costs are incurred at the field collection stage or the laboratory analysis stages. While costs of the former accrue per sample, costs of the latter accumulate per organism. Fewer sampling events may not reduce overall costs if those events collect a greater number of organisms in need of analysis relative to alternative sampling approaches (Trebitz et al. 2010). For surveys such as those recommended here, the need to cover large areas (i.e., large port systems) will require sampling from as many sites as possible to ensure sensitive detection of even heterogeneously distributed rare species. While this will obviously lead to accumulating per sample costs, our recommendation to focus on intensive assessment of relatively few ports will reduce fixed costs per site as much as possible, and also allow cost efficiencies related to post-field sample analysis, particularly if such analyses can be coordinated among surveys.

Nevertheless, previous experience with intensive coastal surveys to assess non-native diversity suggests that the costs of the proposed surveillance program, if pursued aggressively on the time scale recommended, will almost certainly require an investment of tens of millions of dollars. Cohen (2004) provided a thorough assessment of costs associated with long-term (10-year) surveillance for a single large port (Puget Sound). While some of the costs detailed in that prospectus could be avoided, we anticipate adding others (including both greater intensification of sampling and incorporation of genetic analyses) that could significantly elevate expenses. Inglis et al. (2006a) provided a statistically informed estimate of costs associated with obtaining specific detection likelihoods for early-stage incursions of seven high-risk target species. However, that estimate of approximately $1.5 million for a single harbor did not consider

establishing a long-term surveillance program over the decade-long time scale considered in the Cohen study, nor did it include incorporating regular genetic sample analysis. Ultimately, we find it reasonable to suppose that a recipient system surveillance program for a single port over 10 years could require $5 million, resulting in an overall cost of surveys for three to six ports in the $15- to $30-million dollar range.

Table 6-1. Costs associated with nuisance species surveillance programs in terrestrial and aquatic systems

Study (citation)	Estimate cost ($U.S.)	Description
Sliwa et al. (2009b)	$6.16	Realized total cost for baseline survey of 41 Australian ports, 1995-2004; 12 target species
Inglis et al. (2006a)	$1.48	Estimated cost for Whangerei Harbor, New Zealand; based on 90% detection likelihood for detection of 1.5 hectare incursions of 7 target species
Inglis et al. (2006a)	$11.80	Estimated cost for one-time nationwide surveillance for 7 target species in 8 target ports
Ganev and Braithwaite (2003)	$2.02	Estimated cost for 95% probability detection of plum pox virus infestation in New Zealand (inspection of 69,500 trees nationwide)
Stephenson et al. (2003)	$26.56	Estimated cost for 99% probability detection of 12 exotic plant pests in New Zealand (nationwide surveillance)
Cohen (2004)	$5.40	Estimated cost for 10-year marine invasive species surveillance program in Puget Sound; approximately 60 sampling sites

7. SELECTION OF STUDY TARGETS

The following sections discuss important aspects of selecting study targets, including the selection of study taxa and target ports.

7.1 Selection of Study Taxa

Selecting appropriate target taxa will be critically important for experimental approaches. There are two general approaches that might be adopted to taxon selection based on biological criteria: taxa could be chosen because they present worst-case scenarios for invasion or because they are considered "representative" invaders. The former criterion would be used in models aimed at identifying the most precautionary limits on the risk-release relationship, whereas the latter would presumably support development of models that could be adapted across multiple species. We support the former approach for multiple reasons. First, it is not entirely clear that representative invaders exist. Therefore, determining risk-release relationships for worst-case invaders seems more feasible than determining a truly generalizable relationship. Second, the ultimate goal of informing ballast water regulation is well aligned with the decision to parameterize models based on worst cases, as it will presumably help decision-makers determine the most conservative boundaries for future discharge standards. Finally, there is substantial literature aimed at predicting high-risk ANS that provides guidance for selecting worst-case invaders; it is not clear that similar guidance exists on selecting "representative" invaders. These recommendations to explore worst-case scenarios are thus based on pragmatic considerations related to the feasibility of the research effort and the likelihood of developing models that are most useful to managers. Adopting precautionary policy measures is a risk management decision and ultimately outside the purview of the risk assessment research effort being described here, and we are not implying that prioritizing research efforts aimed at assessing worst-case scenarios should drive policy or management. Indeed, other efforts described here, particularly the non-target approaches detailed in sections on ballast water and recipient port surveillance, will expand understanding of the risk-release relationship beyond organisms that have been considered, *a priori*, high risk for establishment. This is critically important not only to prevent model bias, but also to guard against the possibility that those *a priori* assessments of risk fail to accurately capture the true likelihood of establishment for chosen target species.

There are a large number of variables to consider when choosing target study taxa that will maximize invasion success and, consequently, provide a conservative estimate of invasion probability. Perhaps the most obvious are those related to the biology of the organisms. Numerous criteria have been used to determine good candidate species for invasion studies, nuisance species watch lists, and testing the viability of organisms following treatment by ballast water technologies. Important criteria for selecting species that might represent the worst-case scenario include capacity for asexual reproduction, high fecundity (specifically fast-growing in the establishment stage), trophic or habitat generalism, and tolerance of a range of environmental conditions (Kolar and Lodge 2001, 2002, Keller et al. 2007, Bailey et al. 2009, Wonham et al. 2013). For instance, multiple experimental and modeling studies have focused on aquatic taxa that exhibit facultative parthenogenesis or diapausing life history stages (Wonham et al. 2005, Drake et al. 2006, Bailey et al. 2009).

While biological characteristics predisposing taxa to invasiveness are important, other factors will also determine the most appropriate targets for study. For instance, taxa that are conspicuous and readily identifiable will be make shipboard and port surveillance much easier. Also important is the ability to prevent accidental release or quarantine of experimental organisms. Given the priorities of the proposed research, study taxa will ideally be those for which non-native establishment can be, with some reasonable confidence, attributed to ballast water—again, this consideration is particularly important in surveillance programs. For experimental approaches, study taxa should be amenable to laboratory manipulation. The degree to which potential target taxa are suitable for the kinds of experimental study proposed here might be unknown in many cases. Therefore, as noted in Section 4, it is important that protocols be developed that allow experimental populations to survive at relevant densities for relevant time frames. Without such controls, it will be very difficult to assess relationships between propagule supply and establishment. Finally, given recommendations to pursue genetic methods for surveillance, target taxa should ideally be those for which molecular probes have already been or could easily be developed (i.e., those for which substantial genetic information already exists).

Clearly, selecting multiple study taxa is desirable, not only because it may allow characterization of a range of different life histories and reproductive modes, but also because it insures against failure if studying any particular species does not yield the desired results (Carlton et al. 2011, Wonham et al. 2013). Also important will be selecting at least some study taxa that will be adopted for both experimental and descriptive (surveillance) approaches. Ideally, data from descriptive studies would substantiate experimental parameterization of risk-release models.

7.2 Selection of Target Ports

Comprehensively assessing the non-native biodiversity present in North American coastal systems is, without doubt, beyond the scope of any attainable research effort. In fact, given available estimates of the cost of coastal ANS surveillance programs (see Section 6.3), this sort of intensive effort will likely be possible in only a handful of target systems. Therefore, those targets will need to be strategically chosen to leverage resources and to optimize the likelihood of obtaining usable data in a time frame that meets regulatory needs. Various criteria can be used to identify appropriate targets for surveillance. For instance, (Inglis 2001) prioritized ports for surveillance in New Zealand based on both invasion risk and the ecological, economic, social, and cultural values of native ecosystems (Inglis 2001). For the current research effort, surveillance targets might be chosen to maximize information on the variation in parameters likely to influence ANS establishment rates (e.g., environmental matching between recipient and source systems, propagule pressure from multiple vectors, spatio-temporal variance in environmental quality). Unfortunately, capturing such variation between ports may well be beyond the scope of this effort given the small number of target ports. Furthermore, much of this variation could possibly be captured within target systems, if the surveillance effort is intensive enough. We therefore recommend choosing surveillance targets based more on logistical considerations than on any other criteria.

Ports selected for intensive surveillance should:

1. Have well-described invasion histories, known current and historical patterns of ballast water discharge, and well-understood environmental parameters (i.e., there should be adequate baseline knowledge of the system).

2. Have well-described baseline biodiversity and referenced genetic data on resident taxa (native and non-native).

3. Have substantial variation in shipping history, sources of ballast water, and biogeographic location.

4. Provide substantial logistical support for both ship and port sampling (e.g., have a history of sampling and cooperation with local institutions).

5. Be the same for both port and ship surveillance, for at least some subset of targets.

6. Be promising targets for demonstrating the ancillary benefits of the research effort (see below).

Note that these criteria are largely meant to foster cost-effectiveness; meeting them will leverage resources, minimize the amount of additional work required, and communicate the attractiveness and profitability of the effort to as many stakeholders as possible. Note also that many of these criteria are mutually supporting; for instance, those potential target systems with well-described invasion histories most likely also have established histories of ANS monitoring and cooperation with local institutions. In addition, there are benefits to adopting these criteria that are specific to the efficacy of detecting ANS in the target system. Previous studies have indicated that the cost-efficiency of surveillance efforts should increase with increased knowledge of environmental parameters within the study system (Inglis et al. 2006a, Inglis et al. 2006b, Hoffman et al. 2011). Such knowledge allows improved assessments of habitat suitability so sampling efforts can be prioritized. Similarly, understanding current and historical patterns of ballast water discharge may allow managers to develop risk models that prioritize areas for sampling based on the likelihood of dispersal from initial points of inoculation (Inglis et al. 2006b). While studies have indicated that systemic regular sampling is effective in the absence of knowledge about the sampling area (Rew et al. 2006), detection probabilities are known to be strongly influenced by spatial heterogeneity (Harvey et al. 1999, Thomsen et al. 2012a).Therefore, the ability to identify areas that target species are likely to populate greatly decreases the effort required to search (Hoffman et al. 2011); see below. This suggests that choosing surveillance targets based on the availability of baseline data is more than simply a pragmatic approach; there is also considerable theoretical support, based on best scientific understanding of detection likelihoods for rare and invasive species, for adopting this as a primary criterion for target selection.

8. INCORPORATION OF GENETIC TOOLS

The following sections discuss the incorporation of genetic tools, including targeted detection tools, community profiling tools, and cost considerations.

8.1 Background

Effective surveillance entails extracting accurate and detailed information on the species present in sampled communities. This can be time-consuming and expensive, particularly when a premium is placed on identifying rare species. Increasingly, researchers and managers alike are turning to genetic methods to replace or augment monitoring approaches based on traditional morphological identifications. Genetic tools can improve taxonomic resolution, increase sensitivity of detection, and reduce cost per effort for ANS monitoring tasks (Darling and Blum 2007), and large-scale ANS surveillance efforts have long recognized the potential benefits of incorporating genetic methods (Hayes et al. 2005, Inglis et al. 2006a). Many potentially useful tools have been developed in recent years, ranging from species-specific probes for targeted detections to methods that exploit massively parallel DNA sequencing technologies to conduct rapid community profiling (Bott et al. 2010, Wood et al. 2013); see Appendix 11.2 and Table 11-1. In some cases, such tools have been adopted in decision-making contexts (Jerde et al. 2011, Mahon et al. 2011, Mahon et al. 2013), revealing both their usefulness for management and policy purposes and the need for additional research to increase confidence in the inferences drawn from genetic monitoring programs (Darling and Mahon 2011).

Given the ambitious scale of the proposed research, we recommend incorporating genetic detection methods into both ship and port surveillance efforts. The primary aim is to take advantage of potential efficiencies in applying genetic tools in terms of cost per unit effort. A secondary aim is to aggressively pursue developing these tools in the context of biodiversity monitoring (see Section 9.2). Specifically, we recommend developing targeted DNA-based detection methods based on quantitative polymerase chain reaction (PCR) approaches for any target species selected for surveillance, coupled with - community profiling using next-generation sequencing methods. The increased cost of incorporating such tools, particularly in early stages of surveillance when these technologies are still in development, should be well worth the increased sensitivity of detection and taxonomic resolution. More importantly, we anticipate that as the technology advances and confidence increases in the inferences drawn from genetic data, considerable future costs can be saved by relieving some of the burden associated with traditional morphological identifications. This adaptive approach to survey design allows for such adjustments, based on periodically assessing the detection sensitivity, taxonomic resolution, and cost efficiency of genetic tools relative to traditional approaches.

Overlapping goals of ballast water surveillance are determining the presence and abundance of target species of concern and total species diversity with or without a measure of relative abundance. There are several DNA-based technologies that can be applied to these goals, as discussed below. Regardless of application, however, the starting point for each method is DNA or RNA extraction from ballast water. While nucleic acids can be extracted from pre-identified individual organisms, sorting and identifying hundreds or thousands of taxa under a microscope may be prohibitive for large-scale, intensive surveys, and significantly reduces the

expected cost efficiency of molecular monitoring tools. We therefore recommend developing molecular workflows that begin with extraction from unsorted samples.

Sample preservation prior to nucleic acid extraction is critically important, as sample quality can dramatically impact the quality of DNA or RNA and, ultimately, the quality of genetic data (Stein et al. 2013). The most common preservation techniques involve adding ethanol, although this approach can prove challenging for samples with high biomass (Bainard et al. 2010, Nagy 2010). In such cases, ethanol must be periodically replaced to ensure an adequate final concentration of the preservative and full penetration of tissues. An attractive alternative to preservation in ethanol is flash freezing samples in liquid nitrogen or on dry ice. This approach has proven highly effective (Thiyagarajan et al. 2010), but the requirement for very low-temperature refrigeration of samples until they can be processed can add costs. A variety of commercial and non-commercial methods exist for extracting nucleic acids from bulk samples (Ivanova et al. 2006, Kim et al. 2009, Tan and Yiap 2009, Wang et al. 2012). Recommending any specific approach is beyond the scope of this report; however, we do note that there are several critical considerations in choosing a method. First, the method must be capable of extracting DNA from a variety of taxa in an unbiased way. Approaches that disrupt biological material both physically and chemically are most likely to be effective, particularly given the need to process dense samples and the potential size range of taxa. Second, extraction methods must eliminate, as much as possible, compounds present in the environmental sample that inhibit downstream applications (e.g., PCR inhibitors or other inhibitory compounds) (Schrader et al. 2012). Finally, whatever approach is adopted should be cost-effective and readily standardized so that results of genetic analyses can be compared with confidence across multiple sampling sites and multiple surveys.

8.2 Targeted Detection Tools

Thus far, most molecular monitoring tools applied in ANS surveillance contexts have been targeted detection methods. While an array of technologies exist (Darling and Blum 2007, Bott et al. 2010), PCR-based approaches are by far the most commonly used. In fact, for detecting one or few species of concern in a ballast water sample, both endpoint PCR and quantitative PCR (qPCR) represent attractive cost-effective methods. Though the former has been adopted effectively in some management contexts (Darling and Mahon 2011), a significant shortcoming is the inability to translate endpoint PCR detections into reliable estimates of target species abundance, despite the relatively low expense of this approach. qPCR, in contrast, has the potential to be quantitative and allow assessment of measurement uncertainty (Griffiths et al. 2011), and has the additional benefit of having a secondary level of specificity provided by the internal reporting probe. Studies have shown qPCR to be highly sensitive to very rare templates, and controls can provide a measure of quality control. No sorting or visual examination of samples is necessary, further reducing subjectivity in results (Blume et al. 2010). Conventional sorting and searching for rare specimens is time-consuming and labor-intensive. A moderate level of technical expertise is needed for qPCR, but such expertise is taxon-independent. A morphological sorter requires specific training for each taxon. Portable qPCR instruments enable analyses to be done in the field, though processing large numbers of samples is better done in a laboratory. Digital PCR (dPCR), a more recent technological development, similarly allows both quantitation and scaling for field applications (Pohl and Shih 2004).

Although quantifying target species abundance is possible in principle with approaches like qPCR and dPCR, these approaches have not yet been applied to ANS surveillance. While it is relatively straightforward to estimate the concentration of target DNA template in a PCR reaction, and then apply a dilution factor to obtain the concentration of target DNA in the DNA sample, converting target DNA concentration to calculate the abundance of organisms is complicated in multiple ways. Most challenging is the fact that differences in organellar or nuclear genome copies due to variations in body size and potentially numbers of gene copies per genome may result in considerable variability in DNA content per individual. Inferring the number of individuals per environmental sample will require assessing the mean number of target templates per individual, which can be determined in control reactions with DNA from individual or pools of isolated organisms and using pure DNA standards. It is also possible to spike DNA extractions with known numbers of proxy organisms (Mackie and Geller 2010). The advantage of the latter method is that the chemical background of the qPCR reaction is the same as in the tested samples; however, the PCR kinetics for different species may not be identical. The target organism abundance in the plankton sample is calculated as (organisms in DNA sample)/(proportion of entire sample used in DNA extraction). This value can be converted into a density when the volume of water sampled is known.

Non-PCR-based approaches to targeted detection may also be attractive, in part because they may avoid some of the biases inherent to the PCR approach. Sandwich hybridization is a viable approach for detecting one or few taxa. The method can be automated and performed on simple robots (Harvey et al. 2012), or done by hand. A dip-stick format has been used to detect diatoms contributing to harmful algal blooms, for example (Lopez et al. 2008). A downside is that a high copy locus is necessary for sufficient signal strength; thus, ribosomal RNA has been the target of choice. Extracting RNA is more demanding than DNA extraction and requires using fresh plankton or samples stored in RNA later™, an expensive reagent. Samples for DNA extraction can be stored inexpensively in ethanol. Miniarrays (i.e., 96-768 elements) and microarrays (up to ~70,000 elements) are a cost-effective method to detect target sequences in a DNA or RNA extraction. Prior knowledge of diagnostic sequences is necessary, requiring considerable effort to create sequence databases. Low-density miniarrays can be printed with manual or robotic pin-based devices. High-density arrays can be printed with pin-based robots or synthesized onto chips, which allows for the greatest flexibility in modifying chip design. High-density microarrays contain many more elements than the number of multicellular taxa in ballast water samples; therefore, multiple probes can be designed for each known taxon, providing added levels of confidence in a positive signal. Microarrays can be made in batches. Processing requires a moderately high level of technical proficiency that a competent person with an undergraduate level of education can achieve. A specialized scanner is needed to read chips after hybridization.

8.3 Community Profiling Tools

DNA sequencing provides a high level of certainty in species identification because more nucleotide positions are surveyed than in PCR or hybridization-based approaches. In the latter, only 15-60 bp are interrogated, and mismatches between probe/primer are not detectable (unless sufficient to result in no signal). With DNA sequencing, in contrast, up to ~800 bp can be examined from a single sequencing reaction, and any mismatches to a reference sequence are revealed. Because of these advantages, DNA sequencing is frequently used for ANS

identification, particularly when species-level identification is critical to distinguish high-risk invasive species from natives or other low-risk taxa (Darling and Mahon 2011). These applications of traditional Sanger sequencing technology (see Appendix 10.2) likely have limited value for the proposed research, with the possible exception of applications to confirm specificity of targeted genetic surveillance tools. Far more useful will be incorporating next generation sequencing (NGS) technologies (Appendix 10.2) that can rapidly and inexpensively generate vast quantities of DNA sequence data from environmental samples.

NGS technologies have the potential to revolutionize assessments of biodiversity (Baird and Hajibabaei 2012), and could be applied in this research to determine complete community compositions of environmental samples, including both native and non-native taxa—the latter even at low population densities. In addition to increasing taxonomic resolution of identifications and allowing simultaneous assessment of vastly more individuals than standard approaches could reasonably process, NGS tools are also able to provide access to biota that are currently extremely challenging or impossible to query (e.g., many meiofaunal assemblages) (Bik et al. 2012). These tools have already been used successfully in a range of ecological studies, including general biodiversity assessment, studies of diet, food web composition, and temporal changes in species distributions (Yoccoz 2012), and they have demonstrated in multiple systems that biological community structure differs significantly from estimates based solely on traditional morphological approaches (Bik et al. 2012). Particularly attractive in the current context is "DNA metabarcoding," in which automated NGS technologies can identify a large number of species (typically across wide ranges of taxonomic diversity) from a single bulk environmental sample (Taberlet et al. 2012). Such approaches have been lauded as potential solutions to enduring problems in assessing biodiversity. For instance, despite decades of research, the value of standard bioassessments for identifying the causes of aquatic habitat impairment remains limited by insufficient information—small numbers of individuals often identified to genus or higher. NGS technologies could substantially improve the diagnostic power of such assessments by vastly increasing the amount of information across entire biota (Baird and Hajibabaei 2012). ANS surveillance currently experiences many of the same bottlenecks, and thus is likely to similarly benefit from these advances. Metabarcoding approaches have already been shown to be effective tools for understanding aquatic biodiversity in various systems (Chariton et al. 2010, Thomsen et al. 2012a, Thomsen et al. 2012b).

A number of challenges still exist in regular application of metabarcoding approaches for biodiversity assessment. Generally, NGS approaches face difficulties associated with processing and denoising environmental DNA data, picking operational taxonomic units, identifying sequencing errors, and assigning taxonomies (Bik et al. 2012). More specifically, several issues are of particular concern for applications to ANS surveillance. For one thing, we presently assume that NGS is non-quantitative (i.e., the yield of reads per sequence is not tightly correlated to template concentration or, by extension, organism abundance in the environmental sample) (Baird and Hajibabaei 2012). Bias in PCR is presumed to be a major factor that negates this relationship. NGS could be quantitative if PCR were eliminated from the workflow by directly sequencing extracted DNA (Baird and Hajibabaei 2012, Taberlet et al. 2012) and one recent study (Zhou et al. 2013) indicates that a workflow including mitochondrial enrichment and Illumina sequencing (see Appendix 11.2) could allow researchers to sidestep the PCR amplification requirement and achieve strong correlations between sequencing volume and total biomass, allowing estimates of relative abundance. Identifying appropriate barcoding loci also

remains a concern (Tang et al. 2012), although resolving this issue by adopting multiple loci is greatly facilitated by the rapidly shrinking costs of generating sequence data (Baird and Hajibabaei 2012, Yoccoz 2012). Template for DNA sequencing is usually generated through PCR, and PCR primers designed to complement highly conserved sequences allow successful PCR from most taxa; therefore, no *a priori* sequence information is needed. However, if PCR fails for important species of interest, more background work will be needed to design useful PCR primers (e.g., Thomsen et al. 2012a). Another persistent problem with adopting NGS technologies is associated with detecting species at very low population densities—an application critically important in the context of ANS surveillance. Assessments of the so-called "rare biosphere" are prone to biases associated with sequencing error, which can introduce false positives into metabarcoding datasets (Kunin et al. 2010). Thus, although NGS technology has the potential to be highly sensitive, there is some uncertainty about how to treat low frequency or singleton sequences. For instance, one recent study illustrated that, under certain conditions, metabarcoding approaches using the Roche 454 platform could be highly effective at discriminating species of the invasive plant pathogen *Phytophthora*; however, the potential for biasing errors led the authors to recommend "extreme caution" in treating singleton sequences (Vettraino et al. 2012). This is unfortunate, as other studies have shown that in experimentally manipulated samples, extremely rare target templates do show up as singleton sequences, indicating that such results can contain valuable information (Zhan et al. 2013). Resolving these uncertainties will be necessary before the full potential of NGS tools can be used effectively in ANS surveillance; incorporating genetic tools into the proposed research will represent an important effort toward that end.

Despite these challenges, investing in metabarcoding approaches will substantially enhance the information collected in ANS surveillance efforts. This is likely already true, despite the relative immaturity of the technology (Bik et al. 2012). Moreover, extremely rapid advances in both analytical approaches and the cost-efficiency of data generation virtually guarantee that DNA sequencing analysis will be an indispensable component of future biodiversity assessments. The degree to which future tools fit the needs of ANS surveillance depends in large part on investments made now in method development.

8.4 Cost Considerations

Targeted surveillance. Hayes et al. (2005) investigated the cost and detection efficiency of a PCR-based approach for assessing presence of planktonic larvae of the invasive Asian green mussel *Perna viridis* throughout Trinity Inlet near Cairns, New Zealand (Hayes et al. 2005). Cost estimates for genetic detection ran to approximately $2,300 per day, during which 18 samples could be taken (approximately $125/sample). The cost-efficiency of modeled detection by genetic probe was exceeded in their study only by snorkel transects under conditions of good visibility. Unfortunately, visibility is not good in Trinity Inlet, and detection of plankton tows for larvae by divers under poor visibility or by visual inspection was considerably more expensive. Genetic approaches could thus theoretically provide higher level of detection than the next best realistic option at a fraction of the cost; however, the study also noted that these cost savings were somewhat offset by the narrow window of opportunity to sample during the planktonic larval stages.

NGS surveillance. During 2011 and 2012, researchers associated with the Canadian Aquatic Invasive Species Network (CAISN) conducted risk-based sampling for an ANS surveillance program at 16 ports across the country. Sampling was geographically stratified, with equal coverage of all four aquatic realms (east and west coasts, Arctic, Great Lakes-St. Lawrence River). In each realm, four ports with the highest exposure to ANS originating from foreign ballast discharges were sampled, with two exceptions. Sampling in the Arctic included Steenby Inlet, a region that currently has little vessel traffic but which is slated for large increases in shipping, and sampling in the Great Lakes included Nanticoke, Ontario, because of enormous volumes of domestic (i.e., Laker) traffic. Samples were preserved and processed, initially, for three gene markers – 16S, 18S and COI – using 454 pyrosequencing. The greatest species resolution was found using 18S, and thus all further pyrosequencing used only this gene. Four individuals from different taxonomic groups were used to assess intra-individual variation, with numbers of operational taxonomic units (OTUs) dropping to the maximum expected two at 3% genetic divergence. Genbank was queried to determine the intra-interspecies cutoff for 18S for three groups of concern (ascidians, crustaceans, molluscs), and a level of 2% (which captured >90% of species differences) was ultimately adopted. Total genetic divergence to assess species limits was therefore 5% (3% + 2%). All data (half plate per port; four initial ports were represented by between 684,163 and 877,078 initial sequences) were sequentially processed using: 1) CLOTU software to remove <250BP sequences, homopolymers, and sequences lacking tags; 2) CLOTU clustering using the 5% species-level cutoff; 3) trimming using ClustalW software (trims all sequences to 450BP); 4) Usearch software to remove chimeras; and 5) blasting against Pubmed twice for each sequence to generate a final sequence dataset. Final datasets for these four ports had between 436 and 578 distinct OTUs, which were equated with species. The overall reduction in sequence number following data processing averaged 99.93%. These levels of biodiversity are substantially higher than records based on classical taxonomy for the same ports, suggesting either than molecular methods are superior at finding rare species, or that molecular methods generate spurious results despite data processing to remove errors, or both. The remaining 12 ports are due to be processed early in 2013 in a manner similar to those for the initial four ports.

The sensitivity of pyrosequencing was also explored by spiking individual plankton not reported from North America into samples. Freshwater samples were spiked with marine plankton, and marine samples with freshwater plankton (from South America or China). Similar work was conducted using a single individual larva, whose DNA was serially diluted to obtain subindividual DNA abundances, which were also spiked into samples. These studies used the hypervariable V4 region of the nuclear small subunit ribosomal DNA (V4-nSSU). Spiked individuals were recovered as 'singleton' sequences at biomass fractions as low as 2.3×10^{-5}% using a sequencing depth equivalent 1/24 PicoTiter plate. Eleven of 12 samples spiked with one larva (representing four species) were positive. Results at the subindividual level were less sensitive: five of 12 tested positive at 0.1 individuals per sample, and three of 12 at 0.01 individuals per sample; larger larvae were more likely to be detected than small ones. Results of this work were reported in more detail in Zhan et al. (2013).

For preliminary port analyses, costs were approximately $3,200 per port sample (one-half of a 454 PicoTiter plate), for a total of approximately $12,800 for the four ports studied in 2012 or $51,200 for the full survey (16 ports) (Darling 2013b). These are highly conservative estimates, however; they consider only costs for generating sequence data and not personnel

costs associated with either pre-processing of DNA or data analysis. Furthermore, these preliminary assessments likely underestimate the necessary sampling effort associated with intensive surveillance. Nevertheless, additional cost efficiencies can likely be gained. For instance, similar sequencing coverage could be obtained through the Ion Torrent platform at considerable savings (approximately $200-$300 for 600,000 reads) (Darling 2013a). Given results of spiking experiments, independent samples could possibly be processed at sufficient depth with $1/24^{th}$ of a 454 plate, further reducing costs eight-fold relative to the depth obtained in the port studies. This, of course, does not account for the rapid decreases in costs as the technology advances—currently estimated at 50% decreases every five months (Yoccoz 2012).

9. LOGISTICAL CONSIDERATIONS, COORDINATION, AND BENEFITS OF THE PROPOSED RESEARCH

The following sections discuss building a coordinated research effort, ancilliary benefits of the research program, and possibilities for downscaling and prioritization.

9.1 Building a Coordinated Research Effort

> *"To date, there has been no concerted effort to collect and integrate the data necessary to provide a robust analysis of the risk-release relationship needed to evaluate invasion probability associated with particular ballast water discharge standards." (Carlton et al. 2011)*

Considerable effort has already gone into research that is, in some way, relevant to understanding the risk-release relationship. In particular, a large number of studies have attempted to gather data that could, in principle, be used to parameterize risk-release models (and in fact has been used to do so in the past); this is particularly true of data on propagule supply (e.g., see Appendix 11.1) and patterns of establishment. However, these efforts have almost universally been conducted independently, have made little or no attempt to standardize approaches across different studies, and have focused on different primary questions. In short, they have been conducted without any coordination with the aim of informing the risk-release relationship. In this report, we have attempted to provide guidance for developing an ambitious research program that seeks to efficiently leverage financial and human resources across multiple projects and to generate data most salient for informing risk-release models. As indicated in Figure 9-1, this overall research effort spans a considerable time frame and comprises multiple research components, each with interacting and overlapping aims. Such an effort will require considerable coordination among researchers dedicated to different tasks, working in different regions and on different completion schedules.

The ability to synthesize the results of this effort across regions and tasks will be key to the success of this effort and will need careful consideration early and throughout the process. As indicated in Section 1, management decisions have traditionally been made on a scale beyond even the largest ecosystem scales described in Section 4.3, and well above the scale of ports selected as per Section 7.2. For example, leaving aside certain regional exemptions, Canadian ballast water requirements are national in scope, encompassing three coasts: Atlantic, Pacific, and Arctic. While the United States has regional scale (i.e., state) regulations, it also has national regulations whose scope extends even beyond the Continental United States. The IMO standard is global in scope. Informed risk-tolerance decisions based on products discussed in this document will have to account for inevitable variance in results obtained across taxa and between regions. Policymakers will need scientists to understand how to interpret this variance, even as policymakers also account for many other non-risk considerations (e.g., legal, technical, cost-benefit, and international compatibility considerations) as a part of contemplating regulatory change. If an attempt is to be made to address the problem originally posed to the NRC (to "evaluate the risk of successful establishment of new aquatic nonindigenous species associated with a variety of ballast water discharge limits") on a broad basis, every step taken in the process needs to be a calculated step toward this broad synthesis of results.

The degree of coordination necessary to successfully implement the proposed research varies across components. Experimental and surveillance efforts likely need to be only loosely and informally coordinated, but it is important that surveillance efforts target at least some of the target taxa being studied in experimental systems. Obviously, these efforts must share results, but such exchange could probably occur through standard means (e.g., peer-reviewed publication) instead of more directly coordinated mechanisms. This frees experimental and descriptive approaches to remain geographically detached, which is particularly important given limited availability of facilities capable of supporting mesocosm experimentation (see Section 4.2). It also allows experimental studies to be funded independently, which may be an important overall design consideration. The kinds of experimental studies being proposed could possibly be designed so that they are attractive to certain basic science funding agencies (e.g., National Science Foundation) in a way that large-scale, long-term surveillance efforts could not be.

In contrast, ship and port surveillance efforts need to be very tightly coordinated. In fact, we see no way to accomplish the stated aims of the proposed surveillance research without such coordination, and we see a number of crucial benefits to be gained, as discussed below.

Cost-efficiency. The proposed research is expensive. Every effort should be made to ensure that public funds are expended in the most cost-efficient way possible and directed toward the most profitable approaches to filling critical data gaps. Through formal coordination, a large research effort can be much more efficient than the dispersed and disjointed research activities undertaken in the past. For instance, coordinated planning of sampling strategies will save more money than independent planning of sampling over multiple projects, even if such strategies need to be modified based on specific target study area needs and implemented accordingly. Even more valuable is the possibility of leveraging shared resources for training sampling teams or for analyzing biological samples (either morphological or genetic). Shared resources can also reduce costs in terms of reporting, as developing and maintaining coordinated centralized databases will reduce overhead and enable more efficient data archiving than possible across multiple independent efforts.

Standardization. Standardizing approaches across multiple regional efforts is also critically important to the success of the proposed research, particularly in terms of ship and port surveillance. Lack of such standardization is one of the greatest problems associated with deriving reliable inferences based on previously assembled datasets (Carlton et al. 2011, Wonham et al. 2013). Strong top-down coordination of the proposed research will allow implementation of standardized practices. For instance, both ship and port surveillance programs can implement standard sampling designs and practices including training sampling teams in standardized protocols using shared resources. Biodiversity assessments can also be standardized; to the extent that it is possible, both ship and port sampling efforts should utilize the same identification protocols, including genetic approaches. Adopting standardized approaches should also greatly facilitate quality assurance and quality control of data.

Managerial oversight. Substantial investment of public funds requires substantial oversight to ensure that those funds are expended appropriately and that research efforts are progressing toward accepted pre-established goals. The proposed research is both expensive and long-term, and managerial oversight across the multiple components of the program is critical to assure public accountability. Tight coordination of the entire effort will greatly facilitate such

oversight, particularly in the case of ship and port surveillance efforts (see below). A management team should be assembled in the initial planning stages of the proposed research, and that team will be responsible for appropriate distribution of funds and deployment of human resources, and also for making decisions regarding adjustments to the research plan (see below).

Adaptive research planning. Although initial plans can pull from a great deal of existing research, it is clear that there are still important knowledge gaps regarding the most cost-effective sampling strategies, the best choices of target taxa, the most useful technological innovations for diversity assessments, and many other components of the proposed research. It should be recognized at the outset that while the overall aim of the research program is to collect useful data that can be used to parameterize appropriate risk-release models, the program also needs to teach us HOW best to collect such data. The management team will thus have to use an adaptive approach to research planning. For instance, sampling efforts in the past have adopted various forms of adaptive approaches in designing sampling schemes for biodiversity assessment (Acharya et al. 2000, Coe 2008, Brown et al. 2011). Given gaps in our understanding of the likely distribution of target and other non-native species in both ballast tanks and recipient environments, it will be critical to incorporate growing knowledge of such distributions in selecting units to include in future sampling iterations. The adaptive approach to research planning also allows for flexibility in prioritizing and scaling the research effort based on early successes.

Coordinated funding. With the possible exception of experimental approaches (see below), the proposed research will likely not achieve the desired degree of coordination and oversight without coordinated funding. Single-source funding, particularly for the ship and port surveillance efforts, will ensure coordination among regionally dispersed efforts better than if those efforts were funded independently. Past experience suggests that standardizing sampling plans, cost effectively leveraging resources, appropriate sharing data, and the various other benefits of coordination indicated above will be nearly impossible if we rely solely on the good intentions of independent research teams. This creates obvious challenges given the scale of the proposed research, particularly in the present budgetary climate. Nevertheless, it is a critical consideration toward ensuring the success of the overall effort.

Visibility and public outreach. We have described here plans for a substantial, long-term, and costly research effort with clear intended benefits for environmental protection. This is true not only in terms of the support that such research will give to future attempts to develop more environmentally protective approaches to ballast water management, but also in terms of a number of significant ancillary benefits (see Section 9.2). Coordinating the research components described here—particularly if they receive single-source funding—virtually guarantees that the overall effort will receive more public attention that past independent efforts never received, even in aggregate. As discussed above, this will likely attract an unprecedented level of scrutiny and requirement for oversight and public accountability. Nonetheless, it will also provide unprecedented opportunities to communicate the value and benefits of the research conducted to understand ANS and their spread.

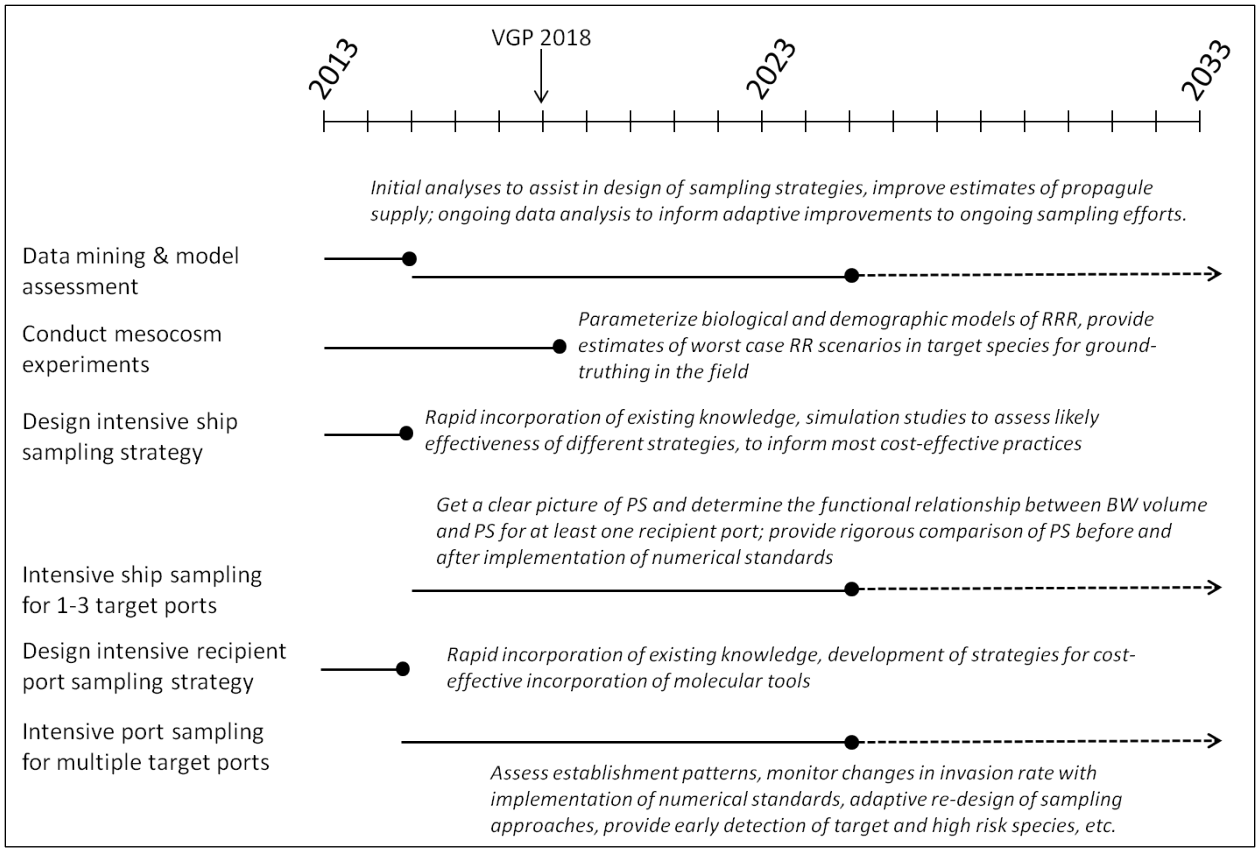

Figure 9-1. Proposed Timeline for Research

Filled circles indicate anticipated deliverables from subprojects; final deliverables anticipated in 2015 include results of long-term port and ship surveillance efforts along with data analysis aimed at establishing most promising general models for assessing the risk-release relationship. Early deliverables include detailed plans for long-term surveillance efforts, informed by preliminary analysis of existing data. Dotted lines indicate recommended continuation of surveillance and data analysis efforts to assess ongoing effectiveness of policy and management decisions. BW = ballast water, PS = , RR = risk-release??, RRR = risk-release relationships??.

9.2 Ancillary Benefits of the Research Program

It will be very difficult to estimate the value of decreasing uncertainty associated with the risk-release relationship, and thus we cannot predict whether improvements in ballast water policy and management would equally improve the effectiveness and efficiency of future regulation. However, while the primary aim of the proposed research is to inform our understanding of the ballast water risk-release relationship, the benefits of the broad, coordinated research effort described in this report would range far beyond the relatively narrow interest of parameterizing models relating propagule supply to establishment risk.

As stated above, while the aim of the proposed research is to understand the ballast water risk-release relationship, better understanding of how propagule supply relates to ANS establishment likelihood will likely decrease uncertainty associated with ANS risk assessment in general, even considering other vectors of introduction (e.g., hull fouling, recreational boating).

Similarly, lessons learned about the risk-release relationship will doubtless increase our knowledge of the mechanisms and patterns of colonization success in aquatic systems, with wide-ranging potential applications in the fields of conservation and restoration biology. Indeed, many of the models most useful for understanding the risk-release relationship are also those frequently applied to understand persistence in natural native populations, particularly those of concern to conservationists (e.g., threatened and endangered species) and resource managers (e.g., fisheries stocks). Coordinated study of experimental and descriptive studies (e.g., linking results of mesocosm population studies with studies of the relationship between propagule supply and establishment rate in ports) will provide unique insights into the validity of inferences drawn from small-scale controlled studies for understanding large-scale natural systems.

The proposed research would also establish a program in target systems for registering changes in propagule supply and invasion rate associated with management and policy changes. As noted previously, each major shift in the regulatory climate shows the impact of policy on the delivery of propagules to recipient systems, with the expectation that reductions in propagule supply will also reduce the rate of establishment of non-native species (Wonham et al. 2013). To a great extent, the global research community has failed to collect the data necessary to test either of these hypotheses on the impact of ballast water exchange (although see Cordell et al. (2009)). Collecting temporal datasets that allow comparison of propagule supplies and establishment rates before and after implementation of recently adopted numerical discharge standards would be of enormous value (Albert et al. 2013). Intensive ship surveillance, if implemented rapidly and before widespread adoption of numerical discharge standards in the United States, could potentially provide reliable estimates of propagule supply reductions resulting from the policy shift. More challenging would be assessing consequent reductions in invasion rates; however, the proposed research effort would make substantial strides toward that goal.

Another important ancillary benefit of this approach will be valuable lessons in designing effective early ANS detection and monitoring at the level of large ports of entry. It is rather remarkable that no standardized approach to long-term surveillance for novel ANS incursions currently exists in the United States (Carlton et al. 2011), and the proposed research would begin to fill this pressing need. Other nations have established standardized surveillance efforts (Hewitt and Martin 2001b, Inglis et al. 2006a, Campbell et al. 2007), and the value of such programs to heighten biosecurity and prevent high-impact aquatic invasions are increasingly being recognized elsewhere in the world. However, even well-established surveillance programs identify the pressing need for additional research to increase detection efficiency and cost-effectiveness (Hayes et al. 2005, Inglis et al. 2006a), and recent efforts continue to assess the detection efficiency of different sampling strategies (Trebitz et al. 2010, Hoffman et al. 2011). However, lessons learned by incorporating these strategies into intensive long-term surveillance approaches such as described here should improve our understanding of the relative cost-effectiveness of various approaches and further refine future efforts. Particularly useful in this regard will be incorporating genetic tools. Applying such tools to monitor changes in metazoan aquatic biodiversity is still in its relative infancy. For instance, next-generation sequencing methods have been used to assess community structure reflecting aquatic habitat quality (Chariton et al. 2010), to characterize marine fish diversity (Thomsen et al. 2012a), and to determine the presence of non-native and other rare species (Zhan et al. 2013); however, all of these studies still could best be considered proofs of concept. Thoroughly exploring the utility of

these tools in ongoing, adaptively designed surveillance efforts would provide opportunities to further refine the application of genetic methods for monitoring ANS and other rare taxa.

More generally, the surveillance program described would also provide a model for establishing a standardized, coordinated monitoring program for understanding the true extent of coastal biodiversity and its response to global change. The past two decades of research on coastal biological invasions has begun to reveal alarming gaps in our knowledge of the historical biogeography of coastal marine taxa (Carlton 2008, Geller et al. 2010), and argues for considerable additional investment in research to better understand the structure of coastal biological communities and temporal patterns of change, particularly those associated with anthropogenic stressors. Aggressively pursuing the recommended research would foster developing and applying novel approaches and technologies (such as genetic tools and novel sampling strategies) that could dramatically enhance the general cost-effectiveness of biodiversity monitoring in coastal and other aquatic systems. Although biodiversity monitoring is a common practice in the United States, particularly as it relates to "bioassessment" of aquatic ecosystem impairment, most programs are highly limited either geographically or taxonomically or both. Standardized national assessments target only a subset of habitats and biota, largely because more extensive monitoring is cost-prohibitive. Lessons learned through developing an effective surveillance program could substantially help develop broader efforts to monitor changes in coastal biodiversity.

The projected cost of the research described here is substantial; however, the potential return on investment is extremely high and extends far beyond the explicit aims of the research program. Such factors are critically important when considering substantial expenditure of public funds, particularly in challenging fiscal climates such as the one we currently face. Improved understanding of the ballast water risk-release relationship, and associated improvements to regulatory structures in the United States and elsewhere, might be sufficient to offset the costs of the proposed research. When one entertains more broadly the ancillary benefits of this research (i.e., generalized knowledge and technological advancements with multiple applications in other fields of conservation and natural resource management), the benefits of a research program such as that described here will ultimately outweigh the costs.

9.3 Possibilities for Down-Scaling and Prioritization

In current and projected funding climates, it may be extremely challenging to justify spending many tens of millions of dollars on a research program aimed at understanding the risk-release relationship associated with ballast water discharge, despite the many direct and indirect benefits outlined above. Although the recommendations outlined here are designed to promote a research program with high likelihood of delivering data relevant to future regulatory needs no available funding vehicle may have the capacity to fund the overall effort. Therefore, opportunities to scale down the proposed research may need to be considered.

The research effort described here could be down-scaled in three ways: (1) conduct the research for a shorter time period; (2) conduct the research in fewer target ports; and (3) conduct less intensive sampling at each target port. If possible, the first of these options should be avoided. The lack of standardized repeated measures of propagule supply and establishment rate is probably the single greatest limitation on existing data and the principle reason that prior

efforts have failed to parameterize a reliable model of the risk-release relationship (Carlton et al. 2011). It is important that any surveillance effort initiated to inform that modeling effort be maintained over a period of at least a decade. Multiple efforts to survey a large number of ports for one or few years would run the risk of providing only marginal gains over available datasets.

Reducing the number of target ports also comes with significant costs. Most troubling is the problem of generalization. Biotic and abiotic differences among recipient systems (e.g., diversity of potential competitors, variance in environmental factors such as salinity and temperature) are likely to substantially influence whether non-native propagules are likely to become established, raising the question of representativeness of ports chosen as study targets. This is a problem in all studies such as this that attempt to inform management and policy at a geographic scale much broader than can be feasibly studied with limited resources. Given a prescribed level of effort based on availability of funding, the highest priority is to design long-term sampling such that one chosen port can be adopted as the target for data collection aimed at informing risk-release models. Remaining funding can be apportioned to efforts in additional ports. Ideally, sufficient funding could be obtained to support full efforts in multiple ports as described above. However, the expense of the complete research program is substantial. Full effort for even a single port (intensive ship and recipient system sampling) over 10 years is likely to cost approximately $10 million, based on estimates outlined in previous sections. This would not include experimental approaches, which (as noted above) could potentially be funded through other independent sources not directly attached to surveillance efforts.

Scaling down efforts within target ports is also a possibility, though this approach would have to be adopted with great caution to prevent diluting the effort so much that statistical robustness of estimates (both propagule supply and establishment rates) is compromised. This suggests that down-scaling should focus less on the number of samples collected and more on the type of data collected. If necessary, sample analysis that focuses on overall description of biotic community structure should be prioritized over efforts to assess density of specific target species. Such descriptions of total biodiversity, with estimates of species density and richness—though more expensive than targeted approaches—will be essential to developing meaningful estimates of overall propagule supply and establishment rate. More detailed assessments of species-specific parameters could then be obtained primarily through complimentary experimental approaches. In the case of extreme limitations to funding, it may be necessary to make additional difficult choices regarding sampling methods, particularly in terms of recipient port sampling. Although the CRIMP protocols have been proven extremely effective, the associated cost per unit effort is relatively high, particularly when compared to passive sampling approaches. As both recommended approaches can be rendered statistically robust and provide opportunities for standardization, it is possible that additional savings could be achieved by adopting less expensive sampling approaches in one or more of the target ports.

In general, all efforts to downscale the proposed research effort will come with costs. However, some costs are greater than others in terms of ultimately achieving the desired endpoint of the research program. Difficult decisions may need to be made in the face of limited available funding, and prioritization of effort along the lines described above may provide options for delivering the most critically important products to support future policy decisions.

10. REFERENCES

16 U.S.C. §§ 4701 et seq. Aquatic Nuisance Prevention and Control. Title 16 U.S. Code, Sec 4701 et seq.

33 U.S.C. §§ 1251 et seq. 2012. Aquatic Nuisance Prevention and Control. 16. Code of Federal Regulations.

77 Federal Register 17254. 2012. Standards for Living Organisms in Ships' Ballast Water Discharged in U.S. Waters. 77.

Acharya, B., G. Bhattarai, A. de Gier, and A. Stein. 2000. Systematic adaptive cluster sampling for the assessment of rare tree species in Nepal. Forest Ecology and Management **137**:65-73.

Albert, R., J. Lishman, and J. Saxena. 2013. Ballast water regulations and the move towards concentration-based numeric discharge limits. Ecological applications : a publication of the Ecological Society of America.

Andersen, M. C. 2005. Potential applications of population viability analysis to risk assessment for invasive species. Human and Ecological Risk Assessment **11**:1083-1095.

Anderson, R. M., and R. M. May. 1991. Infectious diseases of humans: dynamics and control. Oxford University Press, UK.

Bailey, S., M. Deneau, L. Jean, C. Wiley, B. Leung, and H. MacIsaac. 2011. Evaluating Efficacy of an Environmental Policy to Prevent Biological Invasions. Environmental Science & Technology **45**:2554-2561.

Bailey, S., L. Vélez-Espino, O. Johannsson, M. Koops, and C. Wiley. 2009. Estimating establishment probabilities of Cladocera introduced at low density: an evaluation of the proposed ballast water discharge standards. Can J Fish Aquat Sci **66**:261-276.

Bailey, S. A., F. Chan, S. M. Ellis, J. E. Bronnenhuber, J. N. Bradie, and N. Simard. 2012. Risk assessment for ship-mediated introductions of aquatic nonindigenous species to the Great Lakes and freshwater St. Lawrence River. Canadian Science Advisory Secretariat:224.

Bailey, S. A., I. C. Duggan, P. T. Jenkins, and H. J. MacIsaac. 2005. Invertebrate resting stages in residual ballast sediment of transoceanic ships. Canadian Journal of Fisheries and Aquatic Sciences **62**:1090-1103.

Bailey, S. A., I. C. Duggan, C. D. A. van Overdijk, P. T. Jenkins, and H. J. MacIsaac. 2003. Viability of invertebrate diapausing eggs collected from residual ballast sediment. Limnology and Oceanography **48**:1701-1710.

Bainard, L. D., J. N. Klironomos, and M. M. Hart. 2010. Differential effect of sample preservation methods on plant and arbuscular mycorrhizal fungal DNA. Journal of Microbiological Methods **82**:124-130.

Baird, D. J., and M. Hajibabaei. 2012. Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. Molecular Ecology **21**:2039-2044.

Baker, E. 2013. Personal communication with A. Newsom.

Bartell, S. M., and S. K. Nair. 2004. Establishment risks for invasive species. Risk Analysis **24**:833-845.

Bedford, T., and R. Cooke. 2001. Probabilistic risk analysis: foundations and methods. Cambridge University Press.

Beissinger, S. R., and D. R. McCullough. 2002. Population Viability Analysis. University of Chicago Press, Chicago, IL.

Bellwood, D. R., and R. Fisher. 2001. Relative swimming speeds in reef fish larvae. Marine Ecology Progress Series **211**:299-303.

Bik, H. M., D. L. Porazinska, S. Creer, J. G. Caporaso, R. Knight, and W. K. Thomas. 2012. Sequencing our way towards understanding global eukaryotic biodiversity. Trends in Ecology & Evolution **27**:233-243.

Bio-Environmental Services. 1981. The Presence and Implication of Foreign Organisms in Ship Ballast Waters Discharged into the Great Lakes. Georgetown, Ontario.

Blume, L., J. Darling, M. Vazquez, and J. Chandler. 2010. Laboratory audit report: Lodge Laboratory, Department of Biological Sciences, University of Notre Dame.

Bomford, M., F. Kraus, S. C. Barry, and E. Lawrence. 2009. Predicting establishment success for alien reptiles and amphibians: a role for climate matching. Biological Invasions **11**:713-724.

Bott, N. J., K. M. Ophel-Keller, M. T. Sierp, Herdina, K. P. Rowling, A. C. McKay, M. G. K. Loo, J. E. Tanner, and M. R. Deveney. 2010. Toward routine, DNA-based detection methods for marine pests. Biotechnology Advances **28**:706-714.

Boyce, M. S. 1992. Population Viability Analysis. Annual Review of Ecology and Systematics **23**:481-506.

Bradie, J., C. Chivers, and B. Leung. 2013. Importing risk: quantifying the propagule pressure-establishment relationship at the pathway level. Diversity And Distributions.

Briski, E., S. A. Bailey, O. Casas-Monroy, C. DiBacco, I. Kaczmarska, C. Levings, M. L. MacGillivary, C. W. McKindsey, L. E. Nasmith, M. Parenteau, G. E. Piercey, A. Rochon, S. Roy, N. Simard, M. C. Villac, A. M. Weise, and H. J. MacIsaac. 2012a. Relationship between propagule pressure and colonization pressure in invasion ecology: a test with ships' ballast. Proceedings of the Royal Society B-Biological Sciences **279**:2990-2997.

Briski, E., S. A. Bailey, M. E. Cristescu, and H. J. Macisaac. 2010. Efficacy of 'saltwater flushing' in protecting the Great Lakes from biological invasions by invertebrate eggs in ships' ballast sediment. Freshwater Biology **55**:2414-2424.

Briski, E., S. Ghabooli, S. A. Bailey, and H. J. MacIsaac. 2012b. Invasion risk posed by macroinvertebrates transported in ships' ballast tanks. Biological Invasions **14**:1843-1850.

Britton, J. R., J. Pegg, and R. E. Gozlan. 2011. Quantifying imperfect detection in an invasive pest fish and the implications for conservation management. Biological Conservation **144**:2177-2181.

Brown, J. A., M. M. Salehi, M. Moradi, B. Panahbehagh, and D. R. Smith. 2011. Adaptive and unequal probability survey designs for environmental management.

Brown, W. T., M. E. Krasny, and N. Schoch. 2001. Volunteer monitoring of nonindigenous invasive plant species in the Adirondack Park, New York, USA. Natural Areas Journal **21**:189-196.

Burkholder, J., G. Hallegraeff, G. Melia, A. Cohen, H. Bowers, D. Oldach, M. Parrow, M. Sullivan, P. Zimba, E. Allen, C. Kinder, and M. Mallin. 2007. Phytoplankton and bacterial assemblages in ballast water of U.S. military ships as a function of port of origin, voyage time, and ocean exchange practices. HARMFUL ALGAE **6**:486-518.

Campbell, M., B. Gould, and C. Hewitt. 2007. Survey evaluations to assess marine bioinvasions. Marine Pollution Bulletin **55**:360-378.

Cangelosi, A., M. Bain, M. Balcer, D. Reid, E. Reavie, and N. Mays. 2009. Comparison of In-Line and In-Tank Ballast Sampling Methods.

Cangelosi , A., N. Mays, D. Reid, and I. Knight. 2003. Shipboard sampling approaches and recommendations by the Great Lakes Ballast Technology Demonstration Project. Page 15 *in* 1st International Workshop on Guidelines and Standards for Ballast Water Sampling. IMO, Rio de Janeiro, Brazil.

Cangelosi, A., T. Schwerdt, T. Mangan, N. Mays, and K. Prihoda. 2001. A ballast discharge monitoring system for Great Lakes relevant ships: A guidebook for researchers, ship owners, and agency officials. Washington DC.

Carlton, J. T. 1985. Trans-oceanic and Interoceanic Dispersal of Coastal Marine Organisms - The Biology of Ballast Water. Oceanography and Marine Biology **23**:313-371.

Carlton, J. T. 1996. Pattern, process, and prediction in marine invasion ecology. Biological Conservation **78**:97-106.

Carlton, J. T. 2008. Deep Invasion Ecology and the Assembly of Communities in Historical Time.1-44.

Carlton, J. T., and J. B. Geller. 1993. Ecological Roulette - The Global Transport of Nonindigenous Marine Organisms. Science **261**:78-82.

Carlton, J. T., and J. Hodder. 1995. Biogeography and dispersal of coastal marine organisms: experimental studies on a replica of a 16th-century sailing vessel. Marine Biology **121**:9.

Carlton, J. T., A. Navarret, and R. Mann. 1982. Biology of Ships Ballast Water: The Role of Ballast Water in the Transoceanic Dispersal of Marine Organisms. National Science Foundation, Woods Hole, MA.

Carlton, J. T., G. M. Ruiz, J. E. Byers, A. Cangelosi, F. C. Dobbs, E. D. Grosholz, B. Leung, H. J. MacIsaac, and M. J. Wonham. 2011. Assessing the relationship between propagule pressure and invasion risk in ballast water.

Carpenter, S. R. 1998. The need for large-scale experiments to assess and predict the response of ecosystems to perturbation.

Carpenter, S. R., J. J. Cole, J. F. Kitchcell, and M. L. Pace. 2010. Trophic Cascades. Island Press, Washington, D.C. .

Carpenter, S. R., and J. F. Kitchell. 1988. Consumer Control of Lake Productivity. Bioscience **38**:764-769.

Carver, C. E., and A. L. Mallet. 2000. A preliminary assessment of the risks of ballast water-mediated introduction of non-indigenous phytoplankton and zooplankton species into Nova Scotian waters. Transport Canada.

Carver, C. E., and A. L. Mallet. 2002. An assessment of the risks of ballast water-mediated introduction of non-indigenous phytoplankton and zooplankton species into Atlantic Canadian waters. Transport Canada, Dartmouth, Nova Scotia.

Casas-Monroy, O., S. Roy, and A. Rochon. 2011. Ballast sediment-mediated transport of non-indigenous species of dinoflagellates on the East Coast of Canada. Aquatic Invasions **6**:231-248.

Caswell, H. 2001. Matrix population models. 2nd edition. Sinauer Associates, Sunderland, MA.

Cataldo, D., A. Vinocur, I. O'Farrell, E. Paolucci, V. Leites, and D. Boltovskoy. 2012. The introduced bivalve Limnoperna fortunei boosts Microcystis growth in Salto Grande reservoir (Argentina): evidence from mesocosm experiments. Hydrobiologia **680**:25-38.

Chadwell, T. B., and K. A. M. Engelhardt. 2008. Effects of pre-existing submersed vegetation and propagule pressure on the invasion success of Hydrilla verticillata. Journal Of Applied Ecology **45**:515-523.

Chariton, A. A., L. N. Court, D. M. Hartley, M. J. Colloff, and C. M. Hardy. 2010. Ecological assessment of estuarine sediments by pyrosequencing eukaryotic ribosomal DNA. Frontiers In Ecology And The Environment **8**:233-238.

Cherr, G. 2013. Personal communication with A. Newsom.

Choi, K., W. Kimmerer, G. Smith, G. Ruiz, and K. Lion. 2005. Post-exchange zooplankton in ballast water of ships entering the San Francisco Estuary. Journal of Plankton Research **27**:707-714.

Chu, K. H., P. F. Tam, C. H. Fung, and Q. C. Chen. 1997. A biological survey of ballast water in container ships entering Hong Kong. Hydrobiologia **352**:201-206.

Cockrell, M. L., and C. J. B. Sorte. 2013. Predicting climate-induced changes in population dynamics of invasive species in a marine epibenthic community. Journal of Experimental Marine Biology and Ecology **440**:42-48.

Coe, R. 2008. Designing ecological and biodiversity sampling strategies. World Agroforestry Centre.

Cohen, A. N. 2004. An Exotic Species Detection Program for Puget Sound. Puget Sound Action Team, Olympia, Washington.

Cohen, A. N., and J. T. Carlton. 1998. Accelerating invasion rate in a highly invaded estuary. Science **279**:555-558.

Cole, J. J., S. Findlay, and M. L. Pace. 1988. Bacterial Production in Fresh and Saltwater Ecosystems - A Cross-system Overview. Marine Ecology Progress Series **43**:1-10.

Cooke, R. M. 1991. Experts in uncertainty: opinion and subjective probability in science. Oxford University Press, USA.

Cordell, J., D. Lawrence, N. Ferm, L. Tear, S. Smith, and R. Herwig. 2009. Factors influencing densities of non-indigenous species in the ballast water of ships arriving at ports in Puget Sound, Washington, United States. Aquatic Conservation: Marine and Freshwater Ecosystems **19**:322-343.

Costa, R., D. C. Aldridge, and G. D. Moggridge. 2008. Seasonal variation of zebra mussel susceptibility to molluscicidal agents. Journal Of Applied Ecology **45**:1712-1721.

Crecco, V., T. Savoy, and W. Whitworth. 1986. Effects of Density-Dependent and Climatic Factors on American Shad, Alosa-sapidissima, Recruitment - A predictive Approach. Canadian Journal of Fisheries and Aquatic Sciences **43**:457-463.

Cumming, G. S., D. H. M. Cumming, and C. L. Redman. 2006. Scale mismatches in social-ecological systems: Causes, consequences, and solutions. Ecology and Society **11**.

Darling, J. A. 2013a. Personal communication with D Heath.

Darling, J. A. 2013b. Personal communication with EA Brown.

Darling, J. A. 2013c. Personal communication with J Cordell.

Darling, J. A., and M. J. Blum. 2007. DNA-based methods for monitoring invasive species: a review and prospectus. Biological Invasions **9**:751-765.

Darling, J. A., and A. R. Mahon. 2011. From molecules to management: Adopting DNA-based methods for monitoring biological invasions in aquatic environments. Environmental Research **111**:978-988.

David, M., S. Gollasch, M. Cabrini, M. Perkovic, D. Bošnjak, and D. Virgilio. 2007. Results from the first ballast water sampling study in the Mediterranean Sea – the Port of Koper study. Marine Pollution Bulletin **54**:53-65.

David, M., and M. Perkovič. 2003. Ballast water sampling in the Republic of Slovenia. Page 9 *in* 1st International Workshop on Guidelines and Standards for Ballast Water Sampling. IMO, Rio de Janeiro, Brazil.

David, M., and M. Perkovič. 2004. Ballast water sampling as a critical component of biological invasions risk management. Marine Pollution Bulletin **49**:313-318.

Dennis, B., P. L. Munholland, and J. M. Scott. 1991. Estimation of Growth and Extingction Parameters for Endangered Species. Ecological monographs **61**:115-143.

DiBacco, C., D. Humphrey, L. Nasmith, and C. Levings. 2012. Ballast water transport of non-indigenous zooplankton to Canadian ports. ICES Journal of Marine Science **69**:483-491.

Dickman, M., and F. Z. Zhang. 1999. Mid-ocean exchange of container vessel ballast water. 2: Effects of vessel type in the transport of diatoms and dinoflagellates from Manzanillo, Mexico, to Hong Kong, China. Marine Ecology Progress Series **176**:253-262.

Doblin, M. A., L. C. Popels, K. J. Coyne, D. A. Hutchins, S. C. Cary, and F. C. Dobbs. 2004. Transport of the harmful bloom alga Aureococcus anophagefferens by oceangoing ships and coastal boats. Applied and Environmental Microbiology **70**:6495-6500.

Dodgshun, T. J. 2003. Sampling Ships' Ballast Water: The New Zealand Experience. Page 6 *in* 1st International Workshop on Guidelines and Standards for Ballast Water Sampling. IMO, Rio de Janeiro, Brazil.

Drake, J. 2005. Density-Dependent Demographic Variation Determines Extinction Rate of Experimental Populations. PLoS Biology **3**:e222.

Drake, J., and D. Lodge. 2004. Effects of environmental variation on extinction and establishment. Ecology Letters **7**:26-30.

Drake, J. M., K. L. S. Drury, D. M. Lodge, A. Blukacz, N. D. Yan, and G. Dwyer. 2006. Demographic stochasticity, environmental variability, and windows of invasion risk for Bythotrephes longimanus in North America. Biological Invasions **8**:843-861.

Drake, L., M. Doblin, and F. Dobbs. 2007. Potential microbial bioinvasions via ships' ballast water, sediment, and biofilm. Marine Pollution Bulletin **55**:333-341.

Drake, L., G. Ruiz, B. Galil, T. Mullady, D. Friedmann, and F. Dobbs. 2002. Microbial ecology of ballast water during a transoceanic voyage and the effects of open-ocean exchange. Marine Ecology Progress Series **233**:13-20.

Drake, L. A., A. E. Meyer, R. L. Forsberg, R. E. Baier, M. A. Doblin, S. Heinemann, W. P. Johnson, M. Koch, P. A. Rublee, and F. C. Dobbs. 2005. Potential invasion of microorganisms and pathogens via 'interior hull fouling': biofilms inside ballast water tanks. Biological Invasions **7**:969-982.

Drake, L. A., C. S. Moser, S. H. Robbins-Wamsley, S. C. Riley, T. P. Wier, J. F. Grant, P. R. Herring, and M. R. First. Validation trials of a Shipboard Filter Skid (p3SFS) demonstrate its utility for collecting zooplankton.

Ducruet, C., and T. Notteboom. 2012. The worldwide maritime network of container shipping: spatial structure and regional dynamics. Global Networks-a Journal of Transnational Affairs **12**:395-423.

Duggan, I. C., C. D. A. van Overdijk, S. A. Bailey, P. T. Jenkins, H. Limen, and H. J. MacIsaac. 2005. Invertebrates associated with residual ballast water and sediments of cargo-

carrying ships entering the Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences **62**:2463-2474.

Eldredge, L. 1998. Ballast water sampling research: Hawaii (USA). Pages 73-74 USA Ballast Book 1989-1999: Ballast Research in the United States of America.

Endresen, O., H. L. Behrens, S. Brynestad, A. B. Andersen, and R. Skjong. 2004. Challenges in global ballast water management. Marine Pollution Bulletin **48**:615-623.

Estay, S. A., M. Lima, and F. A. Labra. 2009. Predicting insect pest status under climate change scenarios: combining experimental data and population dynamics modelling. Journal of Applied Entomology **133**:491-499.

Fee, E. J., and R. E. Hecky. 1992. Introduction to the Northwest Ontario Lake Size Series (NOLSS). Canadian Journal of Fisheries and Aquatic Sciences **49**:2434-2444.

Fee, E. J., R. E. Hecky, S. E. M. Kasian, and D. R. Cruikshank. 1996. Effects of lake size, water clarity, and climatic variability on mixing depths in Canadian Shield lakes. Limnology and Oceanography **41**:912-920.

Fey, S. B., and K. L. Cottingham. 2012. Thermal sensitivity predicts the establishment success of nonnative species in a mesocosm warming experiment. Ecology **93**:2313-2320.

First, M. R., E. J. Lemieux, W. B. Hyland, J. F. Grant, C. S. Moser, S. C. Riley, S. H. Robbins-Wamsley, M. K. Steinberg, T. P. Wier, and L. A. Drake. 2012. Validation of a closed-housing filter skid for in-line sampling of aquatic organisms. Journal of Plankton Research **34**:321-331.

First, M. R., S. H. Robbins-Wamsley, S. C. Riley, C. S. Moser, G. E. Smith, M. N. Tamburri, and L. A. Drake. 2013. Stratification of Living Organisms in Ballast Tanks: How Do Organism Concentrations Vary as Ballast Water Is Discharged? Environmental Science & Technology **47**:4442-4448.

Floerl, O., and G. J. Inglis. 2005. Starting the invasion pathway: the interaction between source populations and human transport vectors. Biological Invasions **7**:589-606.

Floerl, O., G. J. Inglis, K. Dey, and A. Smith. 2009. The importance of transport hubs in stepping-stone invasions. Journal Of Applied Ecology **46**:37-45.

Fofonoff, R. G.M., H. A.H., S. B.D., and C. J. 2009. Four centuries of biological invasions in tidal waters of the Chesapeake Bay region. Biological invasions in marine ecosystems, Springer-Verlag, Berlin.

Galil, B., and N. Hülsmann. 1997. Protist transport via ballast water — biological classification of ballast tanks by food web interactions. European Journal of Protistology **33**:244-253.

Ganev, S., and M. Braithwaite. 2003. Resource requirements for national active surveillance programmes of high impact exotic pests in New Zealand. Pages 10-15 *in* S. M. Zydenbos, editor. New Zealand Plant Protection, Vol 56.

Geller, J. B., J. A. Darling, and J. T. Carlton. 2010. Genetic Perspectives on Marine Biological Invasions. Annual Review of Marine Science **2**:367-393.

Gertzen, E., B. Leung, and N. Yan. 2011. Propagule pressure, Allee effects and the probability of establishment of invasive species (Bythotrephes longimanus). Ecosphere **2**:1-17.

Gilbert, D. J. 2002. Comment: Does spawning stock affect recruitment of New England groundfish? Interpreting spawning stock and recruitment data in New England groundfish. Canadian Journal of Fisheries and Aquatic Sciences **59**:191-192.

Godwin, L. S., and L. G. Eldredge. 2001. South Oahu Marine Invasions Shipping Study (SOMISS). State of Hawaii Department of Land and Natural Resources.

Gollasch, S. 2006. Overview on introduced aquatic species in European navigational and adjacent waters. Helgoland Marine Research **60**:84-89.

Gollasch, S., M. David, M. Voigt, E. Dragsund, C. Hewitt, and Y. Fukuyo. 2007. Critical review of the IMO international convention on the management of ships' ballast water and sediments. HARMFUL ALGAE **6**:585-600.

Gollasch, S., J. Lenz, M. Dammer, and H. G. Andres. 2000a. Survival of tropical ballast water organisms during a cruise from the Indian Ocean to the North Sea. Journal of Plankton Research **22**:923-937.

Gollasch, S., H. Rosenthal, H. Botnen, J. Hamer, I. Laing, E. Leppakoski, E. Macdonald, D. Minchin, M. Nauke, S. Olenin, S. Utting, M. Voigt, and I. Wallentinus. 2000b. Fluctuations of zooplankton taxa in ballast water during short-term and long-term ocean-going voyages. International Review of Hydrobiology **85**:597-608.

Gollasch, S., H. Rosenthal, I. Laing, E. Leppäkoski, E. Macdonald, D. Minchin, M. Nauke, S. Olenin, S. Utting, M. Voigt, and I. Wallentinus. 1998. Survival rates of species in ballast water during international voyages. Results of the first intercalibration workshop on ship sampling techniques and the first ocean-going workshops of the European Concerted Action: "Testing Monitoring Systems for Risk Assessment of Harmful Introductions by Ships to European Waters".

Gosselin, S., M. Levasseur, and D. Gauthier. 1993. Transport and deballasting of toxic dinoflagellates via ships in the Grande Entrée Lagoon of the Îles-de-la-~adeleine(Gulf of St. Lawrence, Canada).*in* Harmful marine algal blooms - Proceedings of the 6th International Conference on Toxic Marine Phytoplankton.

Grantham, H. S., M. Bode, E. McDonald-Madden, E. T. Game, A. T. Knight, and H. P. Possingham. 2010. Effective conservation planning requires learning and adaptation. Frontiers In Ecology And The Environment **8**:431-437.

Gray, D. K., T. H. Johengen, D. F. Reid, and H. J. MacIsaac. 2007. Efficacy of open-ocean ballast water exchange as a means of preventing invertebrate invasions between freshwater ports. Limnology and Oceanography **52**:2386-2397.

Great Lakes Alliance. 2012. Public comment submitted to the proposed VGP. Comment number EPA-HQ-0141-0724.

Gregg, M., G. Rigby, and G. M. Hallegraeff. 2009. Review of two decades of progress in the development of management options for reducing or eradicating phytoplankton, zooplankton and bacteria in ship's ballast water. Aquatic Invasions **4**:521-565.

Griffiths, K. R., D. G. Burke, and K. R. Emslie. 2011. Quantitative polymerase chain reaction: a framework for improving the quality of results and estimating uncertainty of measurement. Analytical Methods **3**:2201-2211.

Hallegraeff, G. M., and C. J. Bolch. 1991. Transport of Toxic Dinoflagellate Cysts Via Ships Ballast Water Marine Pollution Bulletin **22**:27-30.

Hallegraeff, G. M., and C. J. Bolch. 1992. Transport of Diaom and Dinoflagellate Resting Spores in Ships' Ballast Water - Implications for Plankton Biogeography and Aquaculture. Journal of Plankton Research **14**:1067-1084.

Hallegraeff, G. M., C. J. Bolch, J. Bryan, and B. Koerbin. 1990. Microalgal Spores in Ships' Ballast Water - A Danger to Aquaculture.

Hamer, J. P. 2003. Sampling ballast sediments and other challenges. Page 7 *in* 1st International Workshop on Guidelines and Standards for Ballast Water Sampling. IMO, Rio de Janeiro, Brazil.

Hanlon, J. A., R. J. Albert, and R. A. Everett. 2010. Why EPA and U.S. Coast Guard Requested the National Academy of Sciences Study Assessing Numeric Limits for Living Organisms in Ballast Water *in* First Meeting of the Committee on Assessing Numeric Limits for Living Organisms in Ballast Water, Washington DC.

Harvey, J. B. J., J. P. Ryan, R. Mann, C. M. Preston, N. Alvarado, C. A. Scholin, and R. C. Vrijenhoek. 2012. Robotic sampling, in situ monitoring and molecular detection of marine zooplankton. Journal of Experimental Marine Biology and Ecology **413**:60-70.

Harvey, M., M. Gilbert, D. Gauthier, and D. Reid. 1999. A preliminary assessment of risks for the ballast water-mediated introduction of nonindigenous marine organisms in the Estuary and Gulf of St. Lawrence. Pages 1-66. Canadian Technical Report of Fisheries and Aquatic Sciences 2268

Hay, C., S. Handley, T. Dodgshun, M. Taylor, and W. Gibbs. 1997. Cawthron's Ballast Water Research Programme: Final Report 1996-97. New Zealand.

Hayes, K., R. Cannon, K. Neil, and G. Inglis. 2005. Sensitivity and cost considerations for the detection and eradication of marine pests in ports. Marine Pollution Bulletin **50**:823-834.

Hayes, K. R., and S. C. Barry. 2008. Are there any consistent predictors of invasion success? Biological Invasions **10**:483-506.

Herborg, L. M., C. L. Jerde, D. Lodge, G. Ruiz, and H. MacIsaac. 2007. Predicting invasion risk using measures of introduction effort and environmental niche models. Ecological Applications **17**:663-674.

Herborg, L. M., P. O'Hara, and T. W. Therriault. 2009. Forecasting the potential distribution of the invasive tunicate Didemnum vexillum. Journal Of Applied Ecology **46**:64-72.

Hethcote, H. W. 2000. The mathematics of infectious diseases. Siam Review **42**:599-653.

Hewitt, C., and R. Martin. 2001a. Revised protocols for baseline port surveys for introduced marine species: survey design, sampling protocols and specimen handling.1-52.

Hewitt, C. L., M. L. Campbell, R. E. Thresher, R. B. Martin, S. Boyd, B. F. Cohen, D. R. Currie, M. F. Gomon, M. J. Keough, J. A. Lewis, M. M. Lockett, N. Mays, M. A. McArthur, T. D. O'Hara, G. C. B. Poore, D. J. Ross, M. J. Storey, J. E. Watson, and R. S. Wilson. 2004. Introduced and cryptogenic species in Port Phillip Bay, Victoria, Australia. Marine Biology **144**:183-202.

Hewitt, C. L., and R. B. Martin. 2001b. Revised protocols for baseline port surveys for introduced marine species: survey design, sampling protocols, and specimen handling. Technical Report No. 22.

Hines, A., and G. Ruiz. 2000. Biological Invasions of Cold-Water Coastal Ecosystems...Chapters 1-8.1-340.

Hobbie, J. E., and J. J. Cole. 1984. Response of a detrital food web to eutrophication. Bulletin of Marine Science **35**:357-363.

Hoff, M. H. 2004a. Biotic and abiotic factors related to Lake Herring recruitment in the Wisconsin waters of Lake Superior, 1984-1998. Journal of Great Lakes Research **30**:423-433.

Hoff, M. H. 2004b. Biotic and abiotic factors related to rainbow smelt recruitment in the Wisconsin waters of Lake Superior, 1978-1997. Journal of Great Lakes Research **30**:414-422.

Hoff, M. H., M. W. Meyer, J. Van Stappen, and T. W. Fratt. 2004. Relationships between bald eagle productivity and dynamics of fish populations and fisheries in the Wisconsin waters of Lake Superior, 1983-1999. Journal of Great Lakes Research **30**:434-442.

Hoffman, J. C., J. R. Kelly, A. S. Trebitz, G. S. Peterson, C. W. West, and D. Jackson. 2011. Effort and potential efficiencies for aquatic non-native species early detection. Canadian Journal of Fisheries and Aquatic Sciences **68**:2064-2079.

Holeck, K., E. Mills, H. Macisaac, M. Dochoda, R. Colautti, and A. Ricciardi. 2004. Bridging troubled waters: biological invasions, transoceanic shipping, and the Laurentian Great Lakes. Bioscience **54**:919-929.

Hulme, P. E. 2009. Trade, transport and trouble: managing invasive species pathways in an era of globalization. Journal Of Applied Ecology **46**:10-18.

Huston, M. A. 2002. In Predicting Species Occurrences: Issues of Accuracy and Scale. Island Press Washington, DC. .

IMO. 2004a. Adoption of the final Act and any instruments, recommendations and resolutions resulting from the work of the Conference. In I. M. Organization. International Maritime Organization, London, UK.

IMO. 2004b. International convention for the control and management of ship's ballast water and sediments. International Maritime Organization.

IMO. 2008. Guidelines for Ballast Water Sampling (G2). MEPC.173(58). International Maritime Organization.

Inglis, G., H. Hurren, N. Gust, J. Oldman, I. Fitridge, O. Floerl, and B. Hayden. 2006a. Surveillance design for early detection of unwanted exotic marine organisms in New Zealand.1-228.

Inglis, G., H. Hurren, J. Oldman, and R. Haskew. 2006b. Using habitat suitability index and particle dispersion models for early detection of marine invaders. Ecological applications : a publication of the Ecological Society of America **16**:1377-1390.

Inglis, G. J. 2001. Criteria for identifying and selecting high value locations and locations at risk of invasion by exotic marine organisms in New Zealand. Research Report for Ministry of Fisheries Research Project ZBS2000/04 Objectives 1 & 2

Inglis, G. J., N. Gust, G. MacKay, and R. Tasker. 2004. Field surveys to determine the extent of distribution and relative abundance of Charybdis japonica in the Whangarei region. Final Research Report for Ministry of Fisheries Research Project ZBS2001/01 Objective 3. Wellington.

Ivanova, N. V., J. R. Dewaard, and P. D. N. Hebert. 2006. An inexpensive, automation-friendly protocol for recovering high-quality DNA. Molecular Ecology Notes **6**:998-1002.

Jerde, C. L., A. R. Mahon, W. L. Chadderton, and D. M. Lodge. 2011. "Sight-unseen" detection of rare aquatic species using environmental DNA. Conservation Letters **4**:150-157.

Kaluza, P., A. Kolzsch, M. T. Gastner, and B. Blasius. 2010. The complex network of global cargo ship movements. Journal of the Royal Society Interface **7**:1093-1103.

Kasyan, V. V. 2010. Holoplankton of ship ballast water in the Port of Vladivostok. Russian Journal of Marine Biology **36**:167-175.

Keller, R. P., J. M. Drake, M. B. Drew, and D. M. Lodge. 2011. Linking environmental conditions and ship movements to estimate invasive species transport across the global shipping network. Diversity And Distributions **17**:93-102.

Keller, R. P., J. M. Drake, and D. M. Lodge. 2007. Fecundity as a basis for risk assessment of nonindigenous freshwater molluscs. Conservation Biology **21**:191-200.

Kelly, J. M. 1992. Transport of Non-native Organisms via Cargo Ship Ballast Water: Characterizing the Science/Policy Interface. University of Washington, Seattle.

Kelly, J. M. 1993. Ballast Water and Sediments as Mechanisms for Unwanted Species Introductions into Washington-state. Journal of Shellfish Research **12**:405-410.

Kim, J., M. Johnson, P. Hill, and B. K. Gale. 2009. Microfluidic sample preparation: cell lysis and nucleic acid purification. Integrative Biology **1**:574-586.

Klein, G., K. MacIntosh, I. Kaczmarska, and J. M. Ehrman. 2010. Diatom survivorship in ballast water during trans-Pacific crossings. Biological Invasions **12**:1031-1044.

Kolar, C. S., and D. M. Lodge. 2001. Progress in invasion biology: predicting invaders. Trends in Ecology & Evolution **16**:199-204.

Kolar, C. S., and D. M. Lodge. 2002. Ecological predictions and risk assessment for alien fishes in North America. Science **298**:1233-1236.

Kölzsch, A., and B. Blasius. 2011. Indications of marine bioinvasion from network theory An analysis of the global cargo ship network. European Physical Journal B **84**:601-612.

Kramer, A. M., and J. M. Drake. 2010. Experimental demonstration of population extinction due to a predator-driven Allee effect. Journal of Animal Ecology **79**:633-639.

Kunin, V., A. Engelbrektson, H. Ochman, and P. Hugenholtz. 2010. Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. Environmental Microbiology **12**:118-123.

Lavoie, D. M., L. D. Smith, and G. M. Ruiz. 1999. The potential for intracoastal transfer of non indigenous species in the ballast water of ships. Estuarine Coastal Shelf Science **48**:13.

Lawrence, D. J., and J. R. Cordell. 2010. Relative contributions of domestic and foreign sourced ballast water to propagule pressure in Puget Sound, Washington, USA. Biological Conservation **143**:700-709.

Lee, H., D. Reusser, and M. Frazier. 2013. Approaches to setting organism-based ballast water discharge standards. Ecological applications : a publication of the Ecological Society of America **23**:301-310.

Lee, H., D. Reusser, M. Frazier, and G. Ruiz. 2010. Density matters: review of approaches to setting organism-based ballast water discharge standards.1-132.

Lemieux, E. J., S. Robbins, K. Burns, S. Ratcliff, and P. Herring. 2008. Evaluation of Representative Sampling for Rare Populations using Microbeads. CG-D-03-09, U.S. Coast Guard Research and Development Center, Groton, CT.

Lenz, J., H.-G. Andres, S. Gollasch, and M. Dammer. 1996. Final Report of German Ballast Water Ship Sampling Programme.

Leung, B., N. Roura-Pascual, S. Bacher, J. Heikkila, L. Brotons, M. A. Burgman, K. Dehnen-Schmutz, F. Essl, P. E. Hulme, D. M. Richardson, D. Sol, and M. Vila. 2012. TEASIng apart alien species risk assessments: a framework for best practices. Ecology Letters **15**:1475-1493.

Levin, S. A. 1992. The Problem of Pattern and Scale in Ecology. Ecology **73**:1943-1967.

Locke, A., D. M. Reid, W. G. Sprules, J. T. Carlton, and H. C. van Leeuwen. 1991. Effectiveness of mid-ocean exchange in controlling freshwater and coastal zooplankton in ballast water. Canadian Technical Report of Fisheries and Aquatic Sciences.

Locke, A., D. M. Reid, H. C. Vanleeuwen, W. G. Sprules, and J. T. Carlton. 1993. Ballast Water Exchange as a Means of Controlling Dispersal of Fresh-water Organisms by Ships. Canadian Journal of Fisheries and Aquatic Sciences **50**:2086-2093.

Lodge, D. M., C. R. Turner, C. L. Jerde, M. A. Barnes, L. Chadderton, S. P. Egan, J. L. Feder, A. R. Mahon, and M. E. Pfrender. 2012. Conservation in a cup of water: estimating

biodiversity and population abundance from environmental DNA. Molecular Ecology **21**:2555-2558.

Lopez, C. B., Q. Dortch, E. B. Jewett, and D. Garrison. 2008. Scientific Assessment of Marine Harmful Algal Blooms. Interagency Working Group on Harmful Algal Blooms, Hypoxia, and Human Health of the Joint Subcommittee on Ocean Science and Technology, Washington, D.C.

Lovell, S. J., and L. A. Drake. 2009. Tiny Stowaways: Analyzing the Economic Benefits of a US Environmental Protection Agency Permit Regulating Ballast Water Discharges. Environmental Management **43**:546-555.

Macdonald, E. M., and R. D. Davidson. 1997. Ballast Water Project. Final Report. Fisheries Research Services Report Number 3/97. Aberdeen.

MacIsaac, H., J. Borbely, J. Muirhead, and P. Graniero. 2004. Backcasting and forecasting biological invasions of inland lakes. Ecological applications : a publication of the Ecological Society of America **14**:773-783.

MacIsaac, H. J., T. C. Robbins, and M. A. Lewis. 2002. "Biological invasions of aquatic habitats in Europe and the Great Lakes". Modeling ships' ballast water as invasion threats to the Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences **59**:1245-1256.

Mackie, J. A., and J. Geller. 2010. Experimental parameters affecting quantitative PCR of Artemia franciscana: a model for a marine zooplanktonic target in natural plankton samples. Limnology and Oceanography-Methods **8**:337-347.

Mahon, A. R., C. L. Jerde, W. L. Chadderton, and D. M. Lodge. 2011. Using environmental DNA to elucidate the Asian carp (genus Hypophthalmichthys) invasion front in the Chicago Area Waterway System. Integrative and Comparative Biology **51**:E85-E85.

Mahon, A. R., C. L. Jerde, M. Galaska, J. L. Bergner, W. L. Chadderton, D. M. Lodge, M. E. Hunter, and L. G. Nico. 2013. Validation of eDNA Surveillance Sensitivity for Detection of Asian Carps in Controlled and Field Experiments. PLoS ONE **8**.

Maranda, L. 2013. Personal communication w A. Newsome.

Martens. 2012. Public comment submitted to the proposed VGP; Comment number EPA-HQ-0141-0449. https://www.regulations.gov.

Matsuzaki, S. S., K. Mabuchi, N. Takamura, M. Nishida, and I. Washitani. 2009. Behavioural and morphological differences between feral and domesticated strains of common carp Cyprinus carpio. Journal of Fish Biology **75**:1206-1220.

McCarthy, M. A., and H. P. Possingham. 2007. Active adaptive management for conservation. Conservation Biology **21**:956-963.

McCarthy, S. A., and F. M. Khambaty. 1994. International Dissemination of Epidemic Vibrio-Cholerae by Cargo Ship Ballast and other Nonpotable Waters. Applied and Environmental Microbiology **60**:2597-2601.

McCollin, T., A. M. Shanks, and J. Dunn. 2007. The efficiency of regional ballast water exchange: Changes in phytoplankton abundance and diversity. HARMFUL ALGAE **6**:531-546.

McGee, S., R. Piorkowski, and G. Ruiz. 2006. Analysis of recent vessel arrivals and ballast water discharge in Alaska: Toward assessing ship-mediated invasion risk. Marine Pollution Bulletin **52**:1634-1645.

Medcof, J. C. 1975. Living marine animals in a ship's ballast water. Proceedings of the National Shellfisheries Association **65**.

Meyerson, L. A., and H. A. Mooney. 2007. Invasive alien species in an era of globalization. Frontiers In Ecology And The Environment **5**:199-208.

Middleton, M. J. 1982. The Oriental Goby, Acanthogobius-Flavimanus (Temminck and Schlegel), an Introduced Fish in the Coastal Waters of New-South-Wales, Australia. Journal of Fish Biology **21**:513-523.

Miller, A. W., T. Huber, S. Minton, and G. M. Ruiz. 2010. Status and Trends of Ballast Water Management in the United States: Third Biennial Report to the United States Coast Guard. National Ballast Information Clearinghouse, Smithsonian Environmental Research Center, Edgewater, Maryland.

Mills, C., and F. Sommer. 1995. Invertebrate introductions in marine habitats: two species of hydromedusae (Cnidaria) native to the Black Sea, Maeotias inexspectata and Blackfordia virginica, invade San Francisco Bay. Marine Biology **122**:279-288.

Mills, E. L., J. H. Leach, J. T. Carlton, and C. L. Secor. 1993. Exotic Species in the Great Lakes - A History of Biotic Crises and Anthropogenic Introductions. Journal of Great Lakes Research **19**:1-54.

Mimura, H., R. Katakura, and H. Ishida. 2005. Changes of microbial populations in a ship's ballast water and sediments on a voyage from Japan to Qatar. Marine Pollution Bulletin **50**:751-757.

Minton, M., E. Verling, A. Miller, and G. Ruiz. 2005. Reducing propagule supply and coastal invasions via ships: effects of emerging strategies. Frontiers In Ecology And The Environment **3**:304-308.

Morris, W. F., P. L. Bloch, B. R. Hudgens, L. C. Moyle, and J. R. Stinchcombe. 2002. Population viability analysis in endangered species recovery plans: Past use and future improvements. Ecological Applications **12**:708-712.

Morris, W. F., M. Groom, D. Doak, P. Kareiva, J. Fieberg, L. Gerber, P. Murphy, and D. Thomson. 1999. A practical handbook for population viability analysis. The Nature Conservancy, Washington, DC.

Murphy, K., D. Ritz, and C. Hewitt. 2002. Heterogeneous zooplankton distribution in a ship's ballast tanks. Journal of Plankton Research **24**:729-734.

Nagy, Z. T. 2010. A hands-on overview of tissue preservation methods for molecular genetic analyses. Organisms Diversity & Evolution **10**:91-105.

Neubert, M. G., and I. M. Parker. 2004. Projecting rates of spread for invasive species. Risk Analysis **24**:817-831.

NRDC. 2012. Public comment submitted to the proposed VGP. Comment number EPA-HQ-0141-0446. National Resources Defense Council.

NWF. 2012. Public comment submitted to the proposed VGP. Comment number EPA-HQ-0141-0723. National Wildlife Federation.

Olenin, S., S. Gollasch, S. Jonusas, and I. Rimkute. 2000. En-route investigations of plankton in ballast water on a ship's voyage from the Baltic Sea to the open Atlantic coast of Europe. International Review of Hydrobiology **85**:577-596.

Pace, M. L. 2001. Getting it right and wrong - Extrapolations across experimental scales.

Perrings, C., K. Dehnen-Schmutz, J. Touza, and M. Williamson. 2005. How to manage biological invasions under globalization. Trends in Ecology & Evolution **20**:212-215.

Pierce, R. W., J. T. Carlton, D. A. Carlton, and J. B. Geller. 1997. Ballast water as a vector for tintinnid transport. Marine Ecology Progress Series **149**:295-297.

Pohl, G., and L. M. Shih. 2004. Principle and applications of digital PCR. Expert Review of Molecular Diagnostics **4**:41-47.

Porter-Whitaker, A. E., J. S. Rehage, S. E. Liston, and W. F. Loftus. 2012. Multiple predator effects and native prey responses to two non-native Everglades cichlids. Ecology of Freshwater Fish **21**:375-385.

Power, A. G. 1996. Competition between viruses in a complex plant-pathogen. Ecology **77**:1004-1010.

Reusser, D., H. Lee, M. Frazier, G. Ruiz, P. Fofonoff, M. Minton, and A. Miller. 2013. Per capita invasion probabilities: an empirical model to predict rates of invasion via ballast water. Ecological Applications **23**:321-330.

Rew, L., B. Maxwell, F. Dougher, and R. Aspinall. 2006. Searching for a Needle in a Haystack: Evaluating Survey Methods for Non-indigenous Plant Species. Biological Invasions **8**:523-539.

Reynolds, P. L., and J. F. Bruno. 2012. Effects of Trophic Skewing of Species Richness on Ecosystem Functioning in a Diverse Marine Community. PLoS ONE **7**.

Ricciardi, A. 2003. Predicting the impacts of an introduced species from its invasion history: an empirical approach applied to zebra mussel invasions. Freshwater Biology **48**:972-981.

Ricciardi, A. 2006. Patterns of invasion in the Laurentian Great Lakes in relation to changes in vector activity. Diversity And Distributions **12**:425-433.

Ricciardi, A. 2013. Ecological Systems Springer New York.

Ricciardi, A., and H. J. MacIsaac. 2000. Recent mass invasion of the North American Great Lakes by Ponto-Caspian species. Trends in Ecology & Evolution **15**:62-65.

Richard, R. V., J. F. Grant, and E. J. Lemieux. 2008. Analysis of ballast water sampling port designs using computational fluid dynamics. CG-D-01-08, U.S. Coast Guard Research and Development Center, Groton, CT.

Rigby, G., G. Hallegraeff, and C. Sutton. 1997. Ballast water heating and sampling trials on the BHP ship M.V. Iron Whyalla in Port Kembla and en-route to Port Hedland.

Rius, M., X. Turon, V. Ordoñez, and M. Pascual. 2012. Tracking Invasion Histories in the Sea: Facing Complex Scenarios Using Multilocus Data. PLoS ONE **7**:e35815.

Rothlisberger, J. D., D. M. Lodge, R. M. Cooke, and D. C. Finnoff. 2010. Future declines of the binational Laurentian Great Lakes fisheries: the importance of environmental and cultural change. Frontiers In Ecology And The Environment **8**:239-244.

Roy, S. 1994. Analyse d'échantillons de sédiments provenant de réservoirs d'eau de ballast.

Ruiz, G., and C. Hewitt. 2002. Invasive Aquatic Species of Europe Distribution, Impact and Management. Kluwer Academic Publishers, Dordrecht.

Ruiz, G. M., and J. T. Carlton. 2003. Invasive Species: Vectors and Management Strategies. Island Press, Washington

Ruiz, G. M., P. W. Fofonoff, J. T. Carlton, M. J. Wonham, and A. H. Hines. 2000. Invasion of coastal marine communities in North America: Apparent patterns, processes, and biases. Annual Review of Ecology and Systematics **31**:481-531.

Ruiz, G. M., and A. H. Hines. 1997. The Risk of Nonindigenous Species Invasion in Prince William Sound Associated with Oil Tanker Traffic and Ballast Water Management: Pilot Study. Regional Citizen's Advisory Committee of Prince William Sound, Valdez, AK.

Ruiz, G. M., and D. F. Reid. 2007. Current State of Understanding about the Effectiveness of Ballast Water Exchange (BWE) in Reducing Aquatic Nonindigenous Species (ANS)

Introductions to the Great Lakes Basin and Chesapeake Bay, USA: Synthesis and Analysis of Existing Information. Technical Memorandum GLERL-142. NOAA.

Schindler, D. W. 1998. Replication versus realism: The need for ecosystem-scale experiments. Ecosystems **1**:323-334.

Schmitz, O. J. 2003. Environmental Standards: Combined Exposures and Their Effects on Human Beings and Their Environment Springer-Verlag, New York.

Schneider, D. C. 2001. The rise of the concept of scale in ecology. Bioscience **51**:545-553.

Schrader, C., A. Schielke, L. Ellerbroek, and R. Johne. 2012. PCR inhibitors - occurrence, properties and removal. Journal of Applied Microbiology **113**:1014-1026.

Seebens, H., M. T. Gastner, and B. Blasius. 2013. The risk of marine bioinvasion caused by global shipping. Ecology Letters **16**:782-790.

Simberloff, D. 2009. The Role of Propagule Pressure in Biological Invasions. Annual Review of Ecology Evolution and Systematics **40**:81-102.

Simkanin, C., I. Davidson, M. Falkner, M. Sytsma, and G. Ruiz. 2009. Intra-coastal ballast water flux and the potential for secondary spread of non-native species on the US West Coast. Marine Pollution Bulletin **58**:366-374.

Sliwa, C., S. Migus, F. McEnnulty, and K. R. Hayes. 2009a. Marine bioinvasions in Australia. Biological Invasions in Marine.

Sliwa, C., S. Migus, F. McEnnulty, and K. R. Hayes. 2009b. Marine Bioinvasions in Australia. Pages 425-437 *in* G. Rilov and J. A. Crooks, editors. Ecological Studies.

Smith, L. D., M. J. Wonham, L. D. McCann, D. M. Reid, J. T. Carlton, and G. M. Ruiz. 1996. Shipping Study II: Biological Invasions by Nonindigenous Species in United States Wate1.s: Quantifying the Role ofBallast Water and Sediments, Parts I and II. Report No. CG-D-02-97. U. S. Coast Guard and U. S. Department of Transportation, Washington, DC.

Smith, L. D., M. J. Wonham, L. D. McCann, G. M. Ruiz, A. H. Hines, and J. T. Carlton. 1999. Invasion pressure to a ballast-flooded estuary and an assessment of inoculant survival. Biological Invasions **1**:67–87.

Smith, S. 2013. Personal communication with A. Newsom.

Stein, E. D., B. P. White, R. D. Mazor, P. E. Miller, and E. M. Pilgrim. 2013. Evaluating Ethanol-based Sample Preservation to Facilitate Use of DNA Barcoding in Routine Freshwater Biomonitoring Programs Using Benthic Macroinvertebrates. PLoS ONE **8**.

Stephenson, B. P., G. S. C. Gill, J. L. Randall, and J. A. Wilson. 2003. Biosecurity approaches to surveillance and response for new plant pest species. Pages 5-9 *in* S. M. Zydenbos, editor. New Zealand Plant Protection, Vol 56.

Subba Rao, D. V., W. G. Sprules, A. Locke, and J. T. Carlton. 1994. Exotic phytoplankton from ships' ballast waters: risk of potential spread to mariculture sites on Canada's east coast.

Sylvester, F., O. Kalaci, B. Leung, A. Lacoursiere-Roussel, C. C. Murray, F. M. Choi, M. A. Bravo, T. W. Therriault, and H. J. MacIsaac. 2011. Hull fouling as an invasion vector: can simple models explain a complex problem? Journal Of Applied Ecology **48**:415-423.

Taberlet, P., E. Coissac, F. Pompanon, C. Brochmann, and E. Willerslev. 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular Ecology **21**:2045-2050.

Tan, S. C., and B. C. Yiap. 2009. DNA, RNA, and Protein Extraction: The Past and The Present. Journal of Biomedicine and Biotechnology.

Tang, C. Q., F. Leasi, U. Obertegger, A. Kieneke, T. G. Barraclough, and D. Fontaneto. 2012. The widely used small subunit 18S rDNA molecule greatly underestimates true diversity in biodiversity surveys of the meiofauna. Proceedings of the National Academy of Sciences of the United States of America **109**:16208-16212.

Taylor, C. M., and A. Hastings. 2004. Finding optimal control strategies for invasive species: a density-structured model for Spartina alterniflora. Journal Of Applied Ecology **41**:1049-1057.

Thiyagarajan, V., S. Lau, Tsoi M., Zhang W., and Q. PY. 2010. Monitoring bacterial biodiversity in surface sediment using terminal restrictuion fragment length polymorphism analysis (T-RFLP): application to coastal environment. Pages 151-163 *in* A. Ishimatsu and H. J. Lie, editors. Coastal Environmental and Ecosystem Issues of the East China Sea. Nagasaki University, Nagasaki, Japan.

Thomsen, P. F., J. Kielgast, L. L. Iversen, P. R. Moller, M. Rasmussen, and E. Willerslev. 2012a. Detection of a Diverse Marine Fish Fauna Using Environmental DNA from Seawater Samples. PLoS ONE **7**.

Thomsen, P. F., J. Kielgast, L. L. Iversen, C. Wiuf, M. Rasmussen, M. T. P. Gilbert, L. Orlando, and E. Willerslev. 2012b. Monitoring endangered freshwater biodiversity using environmental DNA. Molecular Ecology **21**:2565-2573.

Trebitz, A., C. West, J. Hoffman, J. Kelly, G. Peterson, and I. Grigorovich. 2010. Status of non-indigenous benthic invertebrates in the Duluth–Superior Harbor and the role of sampling methods in their detection. Journal of Great Lakes Research **36**:747-756.

Tuomisto, H., K. Ruokolainen, and M. Yli-Halla. 2003. Dispersal, environment, and floristic variation of western Amazonian forests. Science **299**:241-244.

US EPA. 2010. Generic protocol for the verification of ballast water treatment technology. EPA/600/R–10/146, U.S. Environmental Protection Agency Environmental Technology Verification Program.

US EPA. 2013. National Pollutant Discharge Elimination System (NPDES) Vessel General Permit (VGP) for discharges incidental to the normal operation of vessels: fact sheet. U.S. Environmental Protection Agency.

USCG. 2012. Ballast Water Discharge Standard-Final Programmatic Environmental Impact Statement. U.S. Coast Guard, Washington, DC.

Verling, E., G. Ruiz, L. Smith, B. Galil, A. Miller, and K. Murphy. 2005. Supply-side invasion ecology: characterizing propagule pressure in coastal ecosystems. Proceedings of the Royal Society B-Biological Sciences **272**:1249-1257.

Vettraino, A. M., P. Bonants, A. Tomassini, N. Bruni, and A. Vannini. 2012. Pyrosequencing as a tool for the detection of Phytophthora species: error rate and risk of false Molecular Operational Taxonomic Units. Letters in Applied Microbiology **55**:390-396.

Wang, C., and J. Wang. 2011. Spatial pattern of the global shipping network and its hub-and-spoke system. Research in Transportation Economics **32**:54-63.

Wang, Y., M. Hayatsu, and T. Fujii. 2012. Extraction of Bacterial RNA from Soil: Challenges and Solutions. Microbes and Environments **27**:111-121.

Weber, M. J., and M. L. Brown. 2013. Density-Dependence and Environmental Conditions Regulate Recruitment and First-Year Growth of Common Carp in Shallow Lakes. Transactions of the American Fisheries Society **142**:471-482.

WFRC. 2013. USGS Western Fisheries Research Center. Western Fisheries Research Center.

Williams, R. J., F. B. Griffiths, E. J. Vanderwal, and J. Kelly. 1988. Cargo Vessel Ballast Water as a Vector for the Transport of Non-indigenous Marine Species. Estuarine Coastal And Shelf Science **26**:409-420.

Wilson, E. O. 2000. On the future of conservation biology. Conservation Biology **14**:1-3.

Winemiller, K. O. 2005. Life history strategies, population regulation, and implications for fisheries management. Canadian Journal of Fisheries and Aquatic Sciences **62**:872-885.

Wonham, M., J. Byers, E. Grosholz, and B. Leung. 2013. Modeling the relationship between propagule pressure and invasion risk to inform policy and management. Ecological Applications:1-56.

Wonham, M., W. Walton, G. Ruiz, A. Frese, and B. Galil. 2001. Going to the source: role of the invasion pathway in determining potential invaders. Marine Ecology Progress Series **215**.

Wonham, M. J., S. A. Bailey, H. J. MacIsaac, and M. A. Lewis. 2005. Modelling the invasion risk of diapausing organisms transported in ballast sediments. Canadian Journal of Fisheries and Aquatic Sciences **62**:2386-2398.

Wonham, M. J., and J. T. Carlton. 2005. Trends in marine biological invasions at local and regional scales: the Northeast Pacific Ocean as a model system. Biological Invasions **7**:369-392.

Wonham, M. J., J. T. Carlton, G. M. Ruiz, and L. D. Smith. 2000. Fish and ships: relating dispersal frequency to success in biological invasions. Marine Biology **136**:1111-1121.

Wonham, M. J., W. C. Walton, A. M. Frese, and G. M. Ruiz. 1996. Transoceanic transport of ballast water: biological and physical dynamics of ballasted communities and the effectiveness of mid-ocean exchange. U. S. Fish and Wildlife Service, Washington, DC.

Wood, S. A., K. F. Smith, J. C. Banks, L. A. Tremblay, L. Rhodes, D. Mountfort, S. C. Cary, and X. Pochon. 2013. Molecular genetic tools for environmental monitoring of New Zealand's aquatic habitats, past, present and the future. New Zealand Journal of Marine and Freshwater Research **47**:90-119.

Yoccoz, N. G. 2012. The future of environmental DNA in ecology. Molecular Ecology **21**:2031-2038.

Zenetos, A., S. Gofas, C. Morri, A. Rosso, D. Violanti, J. E. G. Raso, M. E. Cinar, A. Almogi-Labin, A. S. Ates, E. Azzurro, E. Ballesteros, C. N. Bianchi, M. Bilecenoglu, M. C. Gambi, A. Giangrande, C. Gravili, O. Hyams-Kaphzan, P. K. Karachle, S. Katsanevakis, L. Lipej, F. Mastrototaro, F. Mineur, M. A. Pancucci-Papadopoulou, A. R. Espla, C. Salas, G. San Martin, A. Sfriso, N. Streftaris, and M. Verlaque. 2012. Alien species in the Mediterranean Sea by 2012. A contribution to the application of European Union's Marine Strategy Framework Directive (MSFD). Part 2. Introduction trends and pathways. Mediterranean Marine Science **13**:328-352.

Zhan, A., M. Hulák, F. Sylvester, X. Huang, A. A. Adebayo, C. L. Abbott, S. J. Adamowicz, D. D. Heath, M. E. Cristescu, and H. J. Macisaac. 2013. High sensitivity of 454 pyrosequencing for detection of rare species in aquatic communities. Methods in Ecology and Evolution:n/a-n/a.

Zhang, X. S., J. Holt, and J. Colvin. 2000. A general model of plant-virus disease infection incorporating vector aggregation. Plant Pathology **49**:435-444.

Zhou, X., Y. Li, S. Liu, Q. Yang, X. Su, L. Zhou, M. Tang, R. Fu, J. Li, and Q. Huang. 2013. Ultra-deep sequencing enables high-fidelity recovery of biodiversity for bulk arthropod samples without PCR amplification. GigaScience **2**.

Zvyagintsev, A. Y., V. V. Ivin, I. A. Kashin, T. Y. Orlova, M. S. Selina, V. V. Kasyan, O. M. Korn, E. S. Kornienko, V. A. Kulikova, I. P. Bezverbnaya, L. V. Zvereva, V. I. Radashevsky, L. S. Belogurova, A. A. Begun, and A. N. Gorodkov. 2009. Acclimation and introduction of hydrobionts ships' ballast water organisms in the Port of Vladivostok. Russian Journal of Marine Biology **35**:41-52.

11. APPENDICES

11.1 Abbreviated Summary of Published Reports on Ballast Water Sampling

The full summary includes the following additional categories: sampling location, origin of ballast water, target species (categorized as ≥50 µm, ≥10 µm and <50 µm, or indicator organisms, bacteria pathogens, or viruses), number of samples, sampling methods (net, water, sediment, wall scrape, or combination), vouchering and preservation of samples, time frame of collections, location of collections, public access to data, PI affiliation and contact, and URL for access to data and/or analytical results. Not all categories were available for all datasets.

Publication	Brief Description
Verling et al. (2005)	Ballast water samples to look at survivorship of zooplankton in ballast tanks on 25 separate voyages, involving seven different ships (either oil tankers or coal carriers) on three different routes.
Ruiz and Hines (1997)	Plankton from 16 domestic and one foreign oil tanker included 69 different taxonomic groups. Three ships had BWE.
Harvey et al. (1999)	Ballast water and sediment sampling of incoming foreign vessels. Looked for biodiversity and species richness of protistan and metazoan taxa. A total of 292 phytoplanktonic and 89 zooplanktonic taxa were identified in ballast water of 94 ships of foreign origin, and a total of 65 protistan taxa in ballast sediments collected in eightof these ships. Sixty percent and 57% of the phytoplanktonic and zooplanktonic species were NIS.
Drake et al. (2002)	Ballast water samples looked at microbial community in exchanged and unexchanged ballast water holds during the journey from Hadera, Israel to Baltimore, USA.
Wonham et al. (2001)	Taxon-specific trends in plankton mortality were quantified during a 16-day trans-Atlantic ballast water voyage. (Same as Wonham et al. 1996, an unpublished report). Zooplankton consisted of 44 taxa in 12 phyla, and phytoplankton consisted of three dinoflagellates and two diatoms. One crab species (n = 6 live juvenile portunids, *Liocarcinus holstatus*) in deballasted cargo hold.
Kelly (1993)	Samples (sediment and ballast water) from six Japanese woodchip carriers arriving at Tacoma and Port Angeles in 1991 yielded 21 species of phytoplankton and protists from incubated sediments, and at least eight orders of organisms in ballast water from three ships.
Lawrence and Cordell (2010)	Zooplankton samples from the ballast tanks of 372 ships arriving to Puget Sound ports. Sampled to determine propagule pressure of high-risk coastal zooplankton delivered in ballast water to Puget Sound, Washington, USA. See Cordell et al. 2009.
Gray et al. (2007)	Samples were taken to assess ballast water exchange on six vessels heading to European ports. Two tanks per ship were sampled, one w/BWE and one unexchanged control. Nine cladoceran, seven copepod, and 26 rotifer species were identified from both treatments.

Publication	Brief Description
Williams et al. (1988)	Plankton and fish in 23 woodchip carriers included 61 species; most common were copepods, molluscs, larvaceans and barnacles. Sediments from nine woodchip carriers from seven Japanese vessels yielded 32 crustaceans and polychaetes.
Locke et al. (1991)	A study to look at the effectiveness of mid-ocean BWE in eliminating live freshwater zooplankton from ships originating in freshwater ports. Zooplankton belonging to 12 phyla were collected from 59 vessels. Most samples were numerically dominated by copepods, cladocerans, or rotifers.
Locke et al. (1993)	Plankton samples from 86 ships included 110 species of zooplankton in 11 phyla, mainly copepods, cladocerans and rotifers; 57 species and at least 50 other taxa of invertebrates were represented in the zooplankton samples. (Includes Locke et al. 1991 data).
Medcof (1975)	Plankton sampled from tow holds on a Japanese ship as well as in Twofold Bay, Australia, included polychaetes, copepods, amphipods, ostracods and chaetognaths. Specific taxa not identified.
Chu et al. (1997)	81 species from eight animal phyla and five protist phyla identified from ballast water in container ships.
Kasyan (2010)	Ballast water from three ships at Port of Vladivostok (Peter the Great Bay, Sea of Japan) from ports of Japan (Sea of Japan and Pacific Ocean) and China (Yellow Sea and Yangtze River). Holoplankton identified were from seven taxonomic groups, among which copepods (subclass Copepoda, 33 species) and cladocerans (subclass Cladocera, five species) dominated.
Murphy et al. (2002)	Study looked at the vertical distribution of zooplankton in ballast water of a single bulk carrier during two commercial voyages. Taxa were not identified to fine taxonomic resolution and were grouped according to similar morphological characteristics.
Mimura et al. (2005)	Ballast water and sediment samples from one ship to look for viable bacteria and virus cells and to monitor changes during voyage and BWE.
McCollin et al. (2007)	Ballast water for one bulk carrier was sampled on 12 occasions before and after an exchange process. A total of 175 phytoplankton taxa were identified.
Drake et al. (2007)	Ballast water, unpumpable water and sediment (collectively known as residuals), and biofilms were sampled. All habitats contained bacteria and viruses.
Klein et al. (2010)	Diatom ballast water samples from vessels (majority bulk carriers) arriving in Vancouver from various ports. Forty-one diatom taxa (29 species) containing epiflourescent chloroplasts were identified in 4 ballast tanks. Forty-six diatom genera and 63 species were identified in the TPV samples. Eighty-four total diatom species were found in ballast water.

Publication	Brief Description
Gosselin et al. (1993)	60 ballast water samples taken from ships docked at Îles-de-la-Madeleine adeleine carried small concentrations of four potentially toxic dinoflagellates, *Alexandrium* spp. and three *Dinophysis* spp.
Casas-Monroy et al. (2011)	Ballast sediment from 65 cargo ships (general cargo, bulk carriers and oil tankers) representing two major categories: ships undergoing continental and transoceanic voyages to look for NIS. Fifty-one taxa of dinoflagellate cysts belonging to 40 genera were identified, of which 14 species were not yet reported from Canadian coasts, including four potentially harmful or toxic species.
Hallegraeff and Bolch (1991)	Sediments from 31 out of 83 mainly Japanese woodchip, wheat and ore carriers arriving in Australia in 1987-89 (including the 12 already mentioned) contained dinoflagellate cysts, with toxic species in four ships
Olenin et al. (2000)	Phyto- and zooplankton in ballast water on one ship was examined to assess the potential for the transport of ANS between the Baltic Sea and the open Atlantic coast of Europe. Sixty-two phytoplankton taxa of five major phytoplankton taxonomic divisions were identified; seven potentially toxic. Twenty-seven zooplankton taxa of six major zooplankton taxonomic divisions were identified.
Carver and Mallet (2000)	Ballast water samples from 34 ships (16 containers, five general cargo, eight bulk, three tankers, 2 'other') arriving at three Nova Scotia ports. Two hundred twenty-six phytoplankton taxa (6% NIS) and 44 zooplankton (2% NIS) were identified.
Zvyagintsev et al. (2009)	Ballast water, ballast sediment, and fouling organisms from two vessels arriving in Vladivostok, Russia. Thirty-seven taxa of microalgae belonging to five orders, including 15 diatom and 14 dinoflagellate taxa. Zooplankton (no access to full publication).
Carver and Mallet (2002)	Ballast water samples from 98 ships (29 tankers, 21 bulk, 17 containers and 31 general cargo) arriving at 15 ports in four Atlantic provinces, 77% having undergone BWE. 503 taxa identified (424 phytoplankton and 79 zooplankton) with 105 taxa (25%) classified as NIS.
Briski et al. (2012a)	Ballast water and sediment samples to investigate the relationship between propagule and colonization pressure for a variety of diverse taxonomic groups. (From same study as DiBacco et al. 2012?)
Gollasch et al. (2000a)	Plankton was monitored in ballast water during cruise from Singapore and Columbo, Sri Lanka to Bremerhaven, Germany. Thirty diatom species and 24 zooplankton taxa were identified.
Middleton (1982)	Identified four fish and reported mysids in ballast water of a domestic bulk carrier.
Wonham et al. (1996)	Plankton samples from one coal carrier from Israel yielded 23 species of dinoflagellates and invertebrates, numerically dominated by copepods, bivalves, polychaetes and gastropods.
Rigby et al. (1997)	Test effectiveness of heat treatment on bulk carrier "Iron Whyalla" dinoflagellates, diatomes, cyanobacteria, in ballast tanks: small copepods and copepod nauplii (98%), chaetognaths (1.5%).

Publication	Brief Description
Lavoie et al. (1999)	Plankton diversity and abundance in the ballast water of a coal carrier at the beginning and end of seven replicate voyages were estimated.
Roy (1994)	Eight of nine sediment samples collected from the ballast tanks of three ships contained resting cysts of *Alexandrium* spp.
Gollasch et al. (2000b)	Phyto- and zooplankton in ballast water on four (a-d) ships was examined to assess on changing numbers of individuals inBW during voyages. Zooplankton taxa identified (a) 10, (b) 14, (c) 30, (d) 10.
Kelly (1992)	Samples (sediment) from six Japanese woodchip carriers arriving at Tacoma and Port Angeles in 1991 yielded 21 species of phytoplankton and protists from incubated sediments and at least eight orders of organisms in ballast water from three ships.
Bailey et al. (2003)	Sediments from ballast tanks on nine transoceanic ships of no ballast on board (NOBOB) status were collected. Seventeen cladoceran, copepod, and rotifer taxa were identified.
Hallegraeff et al. (1990)	Sediment from ballasted cargo holds in 12 Japanese woodchip carriers arriving in Tasmania yielded 56 phytoplankton species, including abundant diatoms in four ships and dinoflagellates cysts in seven ships.
David et al. (2007)	Aquatic organisms, including bacteria, in ballast water from ships intending to discharge in Slovenian Sea, i.e., Port of Koper, Mediterranean Sea.
Galil and Hülsmann (1997)	Cultured ballast water and sediment samples from 17 ships yielded at least 198 hererotrophs (reported as flagellate, pseudopodial, and cilate forms) plus diatoms, cnidarians, tuibellarians, nematodes, rotifers, gastrotrichs, polychaetes, and copepods. Three hundred sixty-two records of living protozoan species.
McCarthy and Khambaty (1994)	Ballast water samples in five of 19 ships yielded *Vibrio cholerae*, which genetic analysis found to be identical to the strain responsible for the 1991 South American cholera epidemic and found in oysters in Mobile Bay, Alabama.
Drake et al. (2005)	Ballast water and biofilm samples from a commercial steamship sampled four times after arriving a destination port and biofilm from five other ships. Results reported: bulk microbial metrics – bacteria density, virus-like particle abundance, and algal pigment concentration. Includes data from Drake et al. (2001).
Burkholder et al. (2007)	Samples of ballast water held for two to 176 days, with 90% of the tanks undergoing ballast exchange with open ocean waters. One hundred phytoplankton species were identified, including 23 potentially harmful taxa. Phytoplankton species included 59 diatoms, 32 dinoflagellates, two cryptophyte flagellates, one cyanobacterium, one raphidophyte flagellate, four other ochrophyte (golden) flagellates, and two colonial green algae.

Publication	Brief Description
Dickman and Zhang (1999)	Plankton samples from container ships, which took on ballast water in Manzanillo, Mexico, and discharged in Hong Kong, China, some with mid-ocean ballast water exchange. Thirty diatom and four dinoflagellate taxa identified in unexchanged group. Forty-seven diatom and six dinoflagellate taxa identified in BWE group.
Duggan et al. (2005)	Residual sediment and ballast water sampled from 38 NOBOB ships entering the Great Lakes. Seven established Great Lakes' NIS found, including some discovered since ballast water exchange was implemented. Forty-eight nematode, 35 copepod, 20 harpacticoid.
Bailey et al. (2005)	Sediments from 69 ballast tanks on 39 transoceanic ships of NOBOB status were collected. Seventy-one distinct taxa were identified, 21 of which were NIS (consisting exclusively of rotifers and cladocerans).
Bio-Environmental Services (1981)	Plankton samples from 46 ships that had ballasted outside the northwest Atlantic included 132 phytoplankton, seven protist and 35 invertebrate species.
Hay et al. (1997)	Plankton and bottom water samples from tanks with foreign ballast water in 50 container ships, bulk carriers and break bulk carriers arriving at Lyttelton and Nelson yielded live phytoplankton in 80% of tanks, dominated by diatoms, heterotrophic flagellates and dinoflagellates, and live invertebrates in 83% of tanks with arthropods, molluscs and annelids occurring most frequently.
Godwin and Eldredge (2001)	Ballast water and sediment samples from 13 vessels. Most common organisms were ciliated protozoans, diatoms, dinoflagellates, nematodes, platyhelminthes, molluscs, annelids, crustaceans, and chaetognathes.
Pierce et al. (1997)	The plankton samples from Carlton & Geller 1993; were reexamined for tintinnids. Fifty-six out of the159 ships contained a total of 33 tintinnid species from 15 genera.
Smith et al. (1999)	Plankton and benthic organisms from surveys and samples of foreign ballast water from 60 commercial vessels arriving at Port of Baltimore, MD and in the Port of Norfolk, VA. Biotoa found included 15 animal and three protist phyla, two plant divisions and cyanobacteria. A minimum of 221 distinctly different taxa were identified; 188 taxa were from plankton samples.
Wonham et al. (2000)	Samples looked specifically for fish as well as reporting fish that were collected from previous studies (Carlton and Geller 1993, Smith et al. 1999). In this study, 28 new reports of fishes comprising 17 taxa from 15 families, including the first ballast water records for 13 fish taxa.
Choi et al. (2005)	Zooplankton in post-exchange ballast water from container and bulk carrier ships entering San Francisco Bay. Thirty-three taxa were identified. Copepods were the most abundant zooplankton.
Subba Rao et al. (1994)	Plankton samples from 86 foreign vessels. A total of 102 taxa belonging to seven groups were recognized. Sixty-nine diatoms and 30 dinoflagellates were identified. There were 21 potentially bloom forming, red tide, and/or toxigenic algal species. (From samples collected by Locke et al. 1991, 1993).

Publication	Brief Description
Macdonald and Davidson (1997)	Samples of ballast water and ballast tank sediments from 128 sampling visits to ships in 10 Scottish ports.
Hines and Ruiz (2000)	Fourteen different nonindigenous species (13 crustaceans and one fish) identified. Reports from a new survey combined with previous surveys. (Contains data from Ruiz & Hines 1997 and Hines et al. 1998 with additional data from 1999)).
Smith et al. (1996)	Plankton net, whole and bottom water samples in 70 ships from foreign ports yielded representatives from 15 animal and three protist phyla, two plant divisions, and cyanobacteria. A minimum of 282 distinctly different taxa were identified. Organisms came from freshwater, brackish water, open-ocean, and coastal high-salinity habitats.
Carlton and Geller (1993)	Ballast water samples from 159 cargo ships. Plankton included 16 animal and three protist phyla, and 3 plant divisions. Three hundred sixty-seven different taxa found in ballast water samples taken in Coos Bay, Oregon from Japanese cargo ships.
Lenz et al. (1996)	One hundred eighty-six ship sampled. In ballast water: diatoms (95 spp.), Chloro-phyceae (18 spp.), dinoflagellates (8 spp.), copepoda (52 spp.), Rotatoria (10 spp.) Sediment: diatoms (18 species), dinoflagellates (three species) and their cysts (16 species) and Chlorophyceae (2 species).
Hallegraeff and Bolch (1992)	A survey of 343 cargo vessels entering 18 Australian ports were sampled by 1990, with sampling continuing through at least 1993.
Cordell et al. (2009)	Same as Lawrence and Cordell 2010.
Carlton et al. (1982)	Plankton sampled from a variety of ships and routes included three protist, 24 invertebrate and one fish species.
Carlton (1985)	No access to publication.
Gollasch et al. (1998)	Examined survival of plankton organisms in ballast water tanks on a 23-day voyage from Asia to Germany.
Eldredge (1998)	Ballast water sampling research: Hawaii (USA).
Doblin et al. (2004)	Sampled from residual water in no-ballast-on-board (NOBOB) bulk carrier ships arriving in the North American Great Lakes. Seven of 18 ballast water tanks in commercial ships contained *Aureococcus anophagefferens* following transit from foreign ports.
DiBacco et al. (2012)	Zooplankton in the ballast water of transoceanic exchanged (TOE), intracoastal exchanged (ICE), and intracoastal unexchanged (ICU) vessels.

11.2 Brief Survey of Available Genetic Technologies

A species of concern may be detected by the production of a specific PCR product that is detected (usually electrophoretically) at the conclusion (endpoint) of the PCR reaction. In **endpoint PCR**, PCR primers are designed to complement known sequences derived from the target taxon and tested to mismatch related taxa. Because the false positives cannot be definitively excluded *a priori*, secondary analysis of PCR products (such as restriction endonuclease digestion) may be used. The potential for false negatives are assessed by control PCR reactions using a series of dilution of the DNA from the target organism in the whole plankton DNA to determine the lower limit of detection. Several target organisms can be simultaneously assayed by multiplexing the PCR reaction; however, unknown patterns of interactions among primers and between primers and a variable plankton DNA template may make it difficult to interpret the results. Multiplex PCR probably should be avoided. Quantifying target organisms is not possible with endpoint PCR unless individual organisms are sorted and tested.

Presence and abundance of target taxa can be determined by **quantitative PCR (qPCR)**, in which the rate of PCR product accumulation is related to the starting abundance of starting template. PCR product accumulation is detected in real time by optical detection of fluorescence in the reaction tube. Fluorescence is produced in a variety of ways, but two methods are most often used. In the first, SYBR Green, a dye that fluoresces when bound to dsDNA, is added to the PCR reaction. As double stranded PCR product accumulates, so does the amount of bound SYBR Green. Specificity of SYBR Green assays are conferred only by the PCR primers - signal nonspecific PCR product cannot be separated. The second method involves exonuclease release of a fluorescent during each cycle of PCR. Commercialized as the Taqman© assay, this method uses a probe that is labeled with a fluorescent dye and a quenching molecule. When dye and quencher are proximate in the intact probe, no fluorescence is produced. During each cycle of PCR, the probe will bind to single stranded PCR product during the anneal step simultaneous with PCR primer binding. In the next step, primer extension, Taq polymerase will encounter and digest the bound probe, thus separating quencher and dye, which will now fluoresce. Total fluorescence is related to the production of PCR product. In the Taqman assay, specificity comes from both PCR primer and the internal probe; therefore, signal from nonspecific products or primer-dimer is unlikely. Taqman assays can be multiplexed if differently labeled probes are used, and the assayed PCR product may be the same (if one pair of PCR primers are used) or different (if different PCR primers are used for each target). However, potential intermolecular interactions of probes, primers, and templates may undermine the experiment. Multiplex reactions are not recommended for this reason.

Digital PCR combines the simplicity of endpoint PCR with the ability to quantify as in qPCR. Template DNA is diluted to the point where nanoliter droplets contain single template molecules. The nanodroplet also contains PCR reagent. Therefore, endpoint PCR is conducted in the nanodroplet as it is conveyed through microfluidic channels that pass through thermally controlled zones (for the steps of PCR) and finally past a detector. The number of positive reactions is a direct measurement of the number of template molecules.

The presence of species of concern can be determined by the positive hybridization of template nucleic acid to species-specific oligonucleotides. Such **hybridization assays** can be

quantitative: more DNA from the target organism produces higher signal as long as the probe concentration is not limiting. Calibration curves are necessary, as in qPCR.

In the simplest form, DNA or RNA extractions can be bound to a physical support, such as nylon or nitrocellulose filters, and labeled probe hybridized. In such **low density arrays**, DNA is bound to the filters with vacuum manifolds, and the result is called a dot-blot or slot-blot depending on the shape of the manifold. Careful selection of the sequence in the probe makes the assay species-specific, with the same caveats as mentioned for endpoint PCR. The probe label can be fluorescent, requiring fluorescence imaging equipment, or radioactive, requiring X-ray film processing or radiation imaging apparatus. If different fluorescent labels are used, hybridizations can be multiplexed.

Conversely and more usefully, DNA or RNA from ballast water extractions can be hybridized to probe on a physical support such as filters or glass. **Moderate to high density arrays** of probes can be bound or directly synthesized on such substrata, allowing interrogation of DNA or RNA for large numbers of taxa. Probes can be prepared from PCR product and spotted onto glass in the hundreds to tens of thousands. Probes can also be synthesized on chips in the many tens of thousands. The number of possible probes far exceeds the number of species likely to be surveyed, allowing for multiple probes per taxon for increased confidence of positive results. In this approach, the ballast water DNA or RNA must be labeled for detection. When RNA is the extracted molecule, a reverse transcription reaction produces a labeled cDNA, which may represent a protein coding or ribosomal transcript. When DNA is the extracted molecule, PCR is used to generate labeled DNA for the hybridization reaction.

Sandwich hybridization uses a secondary probe that allows for signal amplification, for example by horseradish peroxidase. An RNA molecule from the plankton extraction is hybridized to the substrate-bound probe. A second probe containing digoxigenin molecule is then hybridized to the RNA molecule, which is now sandwiched by the two probes. Horseradish peroxidase is bound to secondary probe with an anti-digoxigenin antibody. The peroxidase reaction produces light that is optically detected.

Sanger, or dideoxynucleotide termination, sequencing has been the conventional form of DNA sequencing for three decades. Sanger sequencing is a single primer extension reaction in which synthesis of new DNA strands is terminated by the incorporation of a dideoxynucleotide. The ratio of deoxynucleotides and dideoxynucleotides (ddNTP) is optimized so that a shortened product terminating at every nucleotide position in the template. The identity of the terminal nucleotide is indicated by the fluorescent tag of the ddNTP. These products are sorted by electrophoresis with 1 bp resolution; therefore, the template sequence is read as the sequence of ddNTPs passing by a fluorescence detector in the electrophoresis apparatus (i.e., capillary DNA sequencer). Earlier generations of Sanger sequencing used different apparatus, but the chemistry was similar. A limitation of Sanger sequencing is that only one template can be sequenced in one reaction, and only on one strand of the dsDNA template: superimposed sequences are generally unintelligible. This limitation requires that templates for sequencing are from one specimen, thus ballast water plankton samples would need to be sorted.

Next generation sequencing (NGS). NGS is a catch-all phrase for a variety of methods that are united by sequencing single molecules in a massively parallel fashion. These methods

are sometimes called second generation sequencing because yet newer technologies (third generation) now exist. In each second generation method, using a different instrument or platform, a complex DNA sample, such as a genome, a transcriptome, or an environmental sample (e.g., ballast water DNA) is fragmented to a workable size and adaptors ligated. Alternatively, short fragments can be generated by PCR and adaptor sequence ligated or included in the PCR primers. In the Roche, Illumina, and Ion Torrent platforms, sequences are determined during synthesis of a complementary DNA molecular ("sequencing by synthesis"). Next generation sequencing can also detect many species simultaneously and is not limited to known sequence. Illumina and SOLID sequencing do not provide sufficiently long sequences to discriminate among closely related species. There is no critical threshold for required sequence length, but empirically we know that 60-100 bp is insufficient for many congeners. The Life Technologies Ion Torrent and Roche 454 have maximum read length of 400-700 bp and are therefore more appropriate for sequencing of diagnostic amplicons. The Ion Torrent sequencing with the high capacity 318 chip is about sixfold less expensive than the Roche 454 FLX Titanium chemistry (~$2,000/run vs. $12,000/run, depending on the level of template preparation that is outsourced) and yields fourfold more reads. With the Illumina and SOLID platforms better suited for genome and transcriptome sequencing, and the Ion Torrent less expensive for similar or greater read numbers and length, it is expected that the Roche 454 system is near its marketable end. However, real-world comparisons of sequence error rates are needed before the Roche 454 system is eliminated from consideration.

In the workflow for the **Roche 454 pyrosequencer**, these fragments are bound to microbeads by capture sequence in one adaptor. Microbeads are mixed with oil and PCR regeant to form an emulsion of micelles, each containing one bead. PCR of this emulsion (emulsion PCR, or ePCR) amplifies the single molecule with all products retained on the bead. This step is necessary to increase the signal from subsequent sequencing steps. The beads are then recovered from emulsion and distributed onto a picotiter plate containing up to ~10^6 picowells. Within the 454 GS FLX+ or 454 GS Junior instrument, a repeating cycle of flow of the four nucleotides across the picoplate is monitored for nucleotide incorporation: release of pyrophosphate during nucleotide incorporation is coupled to a light-emitting reaction that is optically detected. Images collected at each nucleotide flow are processed at the conclusion of the instrument run to determine which nucleotide was incorporated in each flow cycle. Homonucleotide runs are determined by the intensity of the light signal. However, accuracy of calling homonucleotide runs decreases with the length of the run and is an important source of error. The 454 method produces the longest read lengths of the second generation sequencing platforms (400-700 bp).

The current generation Life Technologies **Ion Torrent** PGM sequencer uses a workflow similar to the Roche 454 system in sample preparation (adaptor ligation, ePCR, distribution onto a sequencing surface). The significant difference is that sequencing surface is a semiconductor chip containing up to ~$4x10^6$ wells, each individual wired to monitor H^+ release during nucleotide incorporation. Thus, no optical images need collection, and data is collected in real time. The Proton sequencer produces up to 250 x 10^6. Instrument run times are ~4 hours; therefore, multiple runs can be performed per day, unlike the Roche of Illumina (below) platforms that require 10-26 hours/run. The semiconductor manufacturing reduces the cost of sequencing about tenfold relative to Roche and Illumina. Read lengths are currently limited to 300 bp, with 400 bp chemistry scheduled for release in January 2013. Ion Torrent has also announced a shift from ePCR in template preparation for late 2013. ePCR is the primary

limitation in the fragment size that may be processed, and by extension limiting read length. An isothermal fragment amplification process will allow fragments ~800 bp to be sequenced, though read length will be >400<800. Bidirectional reads should cover sizes relevant to PCR products.

In the **Illumina** platform, the fragments are bound to capture probes arrayed on a glass support, or flow cell. The fragments are then amplified using adaptors as priming sites, and products are retained on additional capture probes surrounding the original template to produce a cluster of clonal templates. These are analogous to the beads in the 454 approach. Modified nucleotides flow simultaneously across the population of clusters; each of the four nucleotides are labeled with a different flourescent dye and a quencher. When a nucleotide is incorporated, it releases both dye and quencher, producing specific fluorescence that is optically monitored. As in the 454 method, a large number of images are later analyzed to map clusters and determine the sequence of color emitted from each. The Ilumina platform can produce up to ~$3x10^9$ sequences per instrument run, depending on the instrument, but read lengths are short (<100 bp) and run times are long (40 hours to 14 days).

The Applied Biosystems **SOLID** platform sequences templates by ligation of dinucleotide tags (ditags); successful ligation is monitored by release of fluorescence. By mapping overlapping ditags, each nucleotide is read twice, and homonucleotide runs are not error prone as in 454, Illumina, and Ion Torrent systems. Read length is generally 50-100 bp, and ~$1.4x10^8$ reads/instrument run can be generated. Run times are 12 days.

The latest generation of DNA sequencers are designed to generate very long reads from few molecules. So-called “third generation” sequencing is not currently well adapted for ballast water surveillance, as these technologies are geared toward very long, single molecule reads. Organisms in ballast water are generally detected by sequencing shorter diagnostic markers. However, one can imagine adapting this capacity toward surveillance purposes. For example, PCR products from whole-sample extractions could be concatenated and sequenced in a single run. One can also conceive of xth generation sequencers in which ballast water DNA could be sequenced without PCR, and diversity of organisms analyzed in silico.

The only available platform is the **Pacific Biosciences RS**. As in other NGS approaches, DNA is fragmented if necessary, and adapters are ligated. The adapters are designed to convert linear fragments into circular DNA. DNA polymerase is then bound to the circular DNA and this complex is deposited into a septaliter sized sequencing chamber. Modified nucleotide release fluorescence as they are incorporated into newly synthesized strands, with the chamber illuminated from below. Read lengths may be as long as 20kb, averaging ~4kb, with about 10,000 reads/instrument run. Run time is about 2 hours.

Oxford Nanopore Technologies is developing a strikingly different approach to DNA sequencing which utilizes membrane-bound protein pore through which a long DNA molecule may pass. Solid-state pores are said to be in development. As a DNA molecule passes through the pore, changes in pore conformation specific to each nucleotide are monitored as changes in the current measured across the membrane. Read lengths are determined by run-time. Estimates of sequence yields for typical 5-hour runs are up to 10^7 reads of up to 10kb on the **GridION** 8000 instrument. The system is designed in a modular way such that total capacity will depend

on the number of modules used. A miniature system that fits into the USB drive of a computer called the **MinION** is said to produce 100,000 reads of about 9 kb in 6 hours.

Table 11-1. Comparison of technologies available for DNA-based detection and monitoring.

	Throughput (specimens/analyzed per sample)	Requires known sequence	Produces quantitative data	Requires sorting of specimens	Allows discovery of taxa
Endpoint PCR	Low	√			
qPCR	Low	√	√*		
Digital PCR	Low	√	√*		
Dot/Slot blot	Low	√	√**		
Sandwich hybrization	Low	√			
Microarrays	High	√			
Conventional sequencing	Low			√	√***
Roche 454	High				√
Ion Torrent	High				√
Illumina	High				****
SOLID	High				****
Pacific Bioscience	?				
Oxford Nanopore	?				

*requires complex calibration steps

**semi-quantitative

**cryptic species may be discovered

****short read lengths limit species discovery