A new toxicity test with the Arctic algae *Nitzschia frigida*

by

Rebecca J. Eldridge

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

in

Environment and Geography

Department of Environment and Geography

University of Manitoba

Winnipeg

Abstract

There is a pressing need to understand the impact of emerging contaminants on Arctic ecosystems; however, most standard toxicity tests are for temperate species. The lack of Arctic-specific tests may result in an underestimation of harm to Arctic organisms and contribute to significant uncertainty in risk assessments. As well, of the Arctic species tests that exist there has been no assessment of their reliability and relevance for use in environmental risk assessment. To help address these concerns, this thesis consisted of two phases: first, we conducted a critical review on the current state of Arctic ecotoxicology where reported effects were summarized, methodological reliability and endpoint relevance were assessed, and future testing needs were identified. We developed an objective scoring system and evaluated a total of 48 individual studies capturing 39 tested compounds, 73 unique Arctic test species, and 95 distinct endpoints published from 1975 to 2021. Our analysis shows that of 253 test substance and species combinations scored, 207 (82%) failed to meet at least one critical study criterion that contributes to data reliability for use in risk assessment. Significant data gaps were identified as it related to standardized toxicity testing with Arctic species, diversity of compounds tested with these organisms, and the inclusion of ecologically relevant sublethal and chronic endpoints. In the second phase of the project, we developed a laboratory bioassay for the Arctic diatom *Nitzschia frigida*. We optimized culturing conditions by testing various environmental conditions to determine which combination resulted in the most growth within a 14-day period. Under the resulting optimized conditions, we exposed *N. frigida* to copper, zinc, and 1-methylnaphthalene, with concurrent exposures of the standard, temperate species *Skeletonema costatum* in order to compare sensitivity. The test was consistent and repeatable based on control performance and EC50 results across three trials. *N. frigida* demonstrated greater sensitivity than *S. costatum* to the two metals, but not to 1-methylnaphthalene. Overall, this thesis has identified a need for ongoing improvement in test conduct and reporting in the scientific literature to support effective risk assessments and has provided new tools to inform environmental risk assessments in Arctic regions.

Acknowledgements

I would like to extend my sincere thanks to everyone who worked to ensure the success of this project. To my co-advisors and mentors, Dr. Ben de Jourdan and Dr. Mark Hanson, your guidance and encouragement throughout this process have been invaluable on both a professional and personal level. You have taught me to think critically, provided me with amazing opportunities, and prepared me with the skills I need to excel in my field. Thank you for everything. To my committee members, Dr. CJ Mundy and Dr. Ken Jeffries, your feedback and expertise on this project throughout the last four years was critical to its success, and I sincerely thank you for your contributions. To Julie Anderson, thank you for your support through the critical review by providing indispensable feedback that significantly improved the quality of the study.

Thank you to the funders of this work, Polar Knowledge Canada, the Multi-Partner Research Initiative, and the University of Manitoba Clayton H. Riddell Faculty of Environment, Earth, and Resources. Thank you to Dr. Andy Juhl for providing the algal culture and for sharing expertise in Arctic algal culturing techniques. Thank you to Chris Bridger and the Huntsman Marine Science Centre for providing the resources needed to bring this project to life. The importance of funding Arctic research cannot be understated, and without these contributions the work would not have been possible.

I am humbled by the amount of support I received from my Huntsman family. From counting algal cells for hours and troubleshooting with me when things went wrong to listening and supporting me when I felt discouraged – I couldn't have done it without you. Thank you to Danielle Philibert for always being on call to answer my chemistry and statistics questions; to Cailey Dow, Hannah Gale, Perrine Delompre, Jason Durante, Elizabeth Dowling, and Peter Bartlett for all the time you spent both working on my project and listening to me talk about it; to Sue Fordham for your leadership and wise counsel; to Dave Goodwin, whose advice I have sought more times than I can count; and to Howard Streight, Jamie Carpenter, Trena Hurley, Danny Craig, Brooke Lord, and Angela Rehhorn for getting me through the busiest time of my life with

your laughter and friendship. Thank you to Kenzie Perry for all the late night study dates, the constant reassurance, and getting me through a Winnipeg winter. Thank you to Ellen Fanning for always being understanding, supportive, and insightful. I am incredibly grateful for your mentorship.

Last, but definitely not least, an overwhelming thank you goes to my family for your constant support. To Mama Bear for always accepting a million phone calls, never knowing whether you'll hear tears of excitement, tears of stress, or chemistry questions on the other end; to Dad and Sherry for always making me feel like I can do anything, and for the garage nights making prototypes of whatever I dreamed up; to my siblings for always reminding me that the chaos would be worth it; and to Charlene and Jeff for your constant encouragement. To Taco, Coal, and Dixie: thank you for never leaving me to work on my thesis alone.

To Mat McGlinchey: there are no words to express how thankful I am to have had you by my side this entire time. Thank you for being my biggest cheerleader, my shoulder, and my rock.

Contents

List of Tables

List of Figures

List of Acronyms

ACAP	Arctic Contaminants Action Program
AMAP	Arctic Monitoring and Assessment Programme
AMOP	Arctic and Marine Oil Spill Program
ANOVA	Analysis of variance
ASTM	American Society for Testing and Materials
ASW	Artificial seawater
AWT	Artificially weathered
BOD	Biochemical oxygen demand
BPP	Biomass-specific primary production
CAFF	Conservation of Arctic Flora and Fauna
CAS	Chemical Abstracts Service
CCME	Canadian Council of Ministers of the Environment
CEWAF	Chemically-enhanced water accommodated fraction of crude oil
COOGER	Centre for Offshore Oil, Gas, and Energy Research
COSEWIC	Committee on the Status of Endangered Wildlife in Canada
CRED	Criteria for reporting and evaluating toxicity data
CROSERF	Chemical Response to Oil Spills: Ecological Research Forum
$CuSO_4$	Copper sulphate
CV	Coefficient of variation
DDT	Dichlorodiphenyltrichloroethane
DEET	N,N-diethyl-meta-toluamide
DFO	Department of Fisheries and Oceans
DO	Dissolved oxygen
DOC	Dissolved organic carbon
$EC50_{min}$	Incipient EC50 value
ECCC	Environment and Climate Change Canada
ECIS	European Conference on Information Systems
ECx	Effective concentration to x% of organisms
EPPR	Emergency Prevention, Preparedness and Response
ERA	Environmental risk assessment
EROD	Ethoxyresorufin-O-deethylase
EU	Experimental unit
FR	Fresh (as opposed to weathered [oil])
GC-MS	Gas chromatography-mass spectrometry
GLP	Good Laboratory Practices
HESAW	Harrison’s medium with artificial seawater
HESNW	Harrison’s medium with natural seawater
HMSC	Huntsman Marine Science Centre
HPLC	High-performance liquid chromatography
HSD	Honestly significant difference
IC	Inoculum control
ICE	Interspecies Correlation Estimation
ICx	Inhibition concentration to x% of organisms
IMO	International Maritime Organization

List of Acronyms (continued)

IOSC	International Oil Spill Conference
ISO	International Organization for Standardization
LCx	Lethal concentration to x% of organisms
LOD	Limit of detection
LOEC	Lowest observed effect concentration
NC	Not calculable
NIWA	National Institute of Water and Atmospheric Research
NM	Not measured
NR	Not reported
NOEC	No observed effect concentration
NPP	Net primary production/productivity
NSAID	Non-steroidal anti-inflammatory drugs
NSW	Natural seawater
NWT	Naturally weathered
OECD	Organization for Economic Co-operation and Development
PAH	Polycyclic aromatic hydrocarbon
PAME	Protection of the Arctic Marine Environment
PAR	Photosynthetically active radiation
PCB	Polychlorinated biphenyl
PDMS	Polydimethylsiloxane
PFAS	Perfluoroalkyl and polyfluoroalkyl substances
PFC	Perfluorinated compounds
PNEC	Predicted no-effect concentration
POP	Persistent organic pollutant
PPCP	Pharmaceuticals and personal care products
RPC	Research and Productivity Council
RPM	Revolutions per minute
SDWG	Sustainable Development Working Group
SSD	Species sensitivity distribution
SSRI	Selective serotonin reuptake inhibitors
TPH	Total petroleum hydrocarbons
US EPA	United States Environmental Protection Agency
WAF	Water accommodated fraction of crude oil
WOE	Weight of evidence
WT	Weathered
$ZnSO_4$	Zinc sulphate

1 Introduction

1.1 Overview

The Canadian Arctic is significant ecologically, economically, socially, and culturally. Increasing populations of Northern communities and expanded resource development has brought with it many challenges and potential impacts that must be addressed; therefore, growing research focus to quantify the impact of anthropogenic activities is needed (Krumhansl et al., 2015). However, there is a gap in scientific knowledge regarding toxicity of contaminants in general to Arctic species (Chapman and Riddle, 2005). The response of relatively numerous temperate aquatic species to a vast array of contaminants has been characterized (Chapman and Riddle, 2003; United States Environmental Protection Agency [US EPA], 2019), but critical differences in the structure and function of Arctic and temperate ecosystems prevent these data from being readily transferrable to the well-adapted species of the North without significant uncertainty. The presence of contaminants in Arctic marine ecosystems underline the importance of generating Arctic-specific toxicity data in order to provide accurate and robust insight for environmental risk assessment (ERA) in the region. This study aims to investigate the hypothesis that Arctic algal species may be more sensitive to contaminants than their standard, temperate counterparts. This was investigated by designing and performing toxicity tests with Arctic algae. In addition, a critical review was conducted to describe the current state of knowledge in Arctic ecotoxicological research in order to identify knowledge gaps in the field and recommend future research needs. A summary of the type, number, and quality of studies performed and the data produced are included. Studies were critically analyzed using a scoring rubric to determine data reliability in the context of ERA.

1.2 The Arctic

Researchers have defined the Arctic in numerous ways depending on context (Figure 1.1, Arctic Centre, University of Lapland, n.d.). It is sometimes referred to as only the region within the Arctic Circle, the boundary of which encompasses the entire

region north of approximately 66.5°N latitude (United States Department of the Interior, 2008). This imaginary line perpetually drifts with shifting planetary patterns as it delineates the limit of 24-hour winter darkness and 24-hour summer daylight. Others denote the boundary as northward of regions that experience an average temperature of below 10°C during the warmest month of the year (Arctic Centre, University of Lapland, n.d.). A third commonly applied definition is the boundary of the terrestrial treeline, north of which trees are unable to thrive due to extreme temperatures (Bellamy, 2010). This study aims to describe the relative sensitivity of Arctic algal species to contaminants and therefore refers to the Arctic as any region that can be characterized by the trophic level interactions that are representative of Arctic ecosystems that are discussed in Section 1.3 of this thesis, as opposed to any of the more specific aforementioned definitions. Currently, eight countries have a geographic claim in the Arctic, including Canada, Denmark, Finland, Iceland, Norway, Russia, Sweden, and the United States (Arctic Council, 2015). Canada is home to a large proportion of this, with just under two thirds of its land situated north of the limit of isolated permafrost (Natural Resources Canada, 2017). These multiple stakeholders carry major political and economic power on a global scale and therefore have the ability to drastically alter the Arctic landscape, and hence in part the need to ensure appropriate tools are available to screen for effects.

At present, the Arctic Council is the main driver of multinational efforts to regulate contamination of Arctic ecosystems as a result of development. It is an intergovernmental institution comprised of the eight aforementioned Arctic states and Arctic Indigenous communities whose mission statement is to promote cooperation, coordination, and interaction among the different stakeholders to address issues of sustainable development and environmental protection in the region (Arctic Council, 2020). The Arctic Council operates under a “principle of consensus”, meaning that all eight Arctic Member states must reach consensus before decisions are made. Since its establishment in 1996, six working groups have been formed under the umbrella of the Arctic Council to achieve its mission statement. These include the Arctic Contaminants Action Program (ACAP), the Arctic Monitoring and Assessment Programme (AMAP), Conservation of Arctic Flora and Fauna (CAFF), Emergency Prevention, Preparedness,

and Response (EPPR), Protection of the Arctic Marine Environment (PAME), and the Sustainable Development Working Group (SDWG). These working groups are responsible for conducting scientific studies, providing forums for intergovernmental communication, and making recommendations to the Arctic states. The importance of multinational cooperation in Arctic research cannot be understated, particularly in relation to scientific research and environmental contaminants.

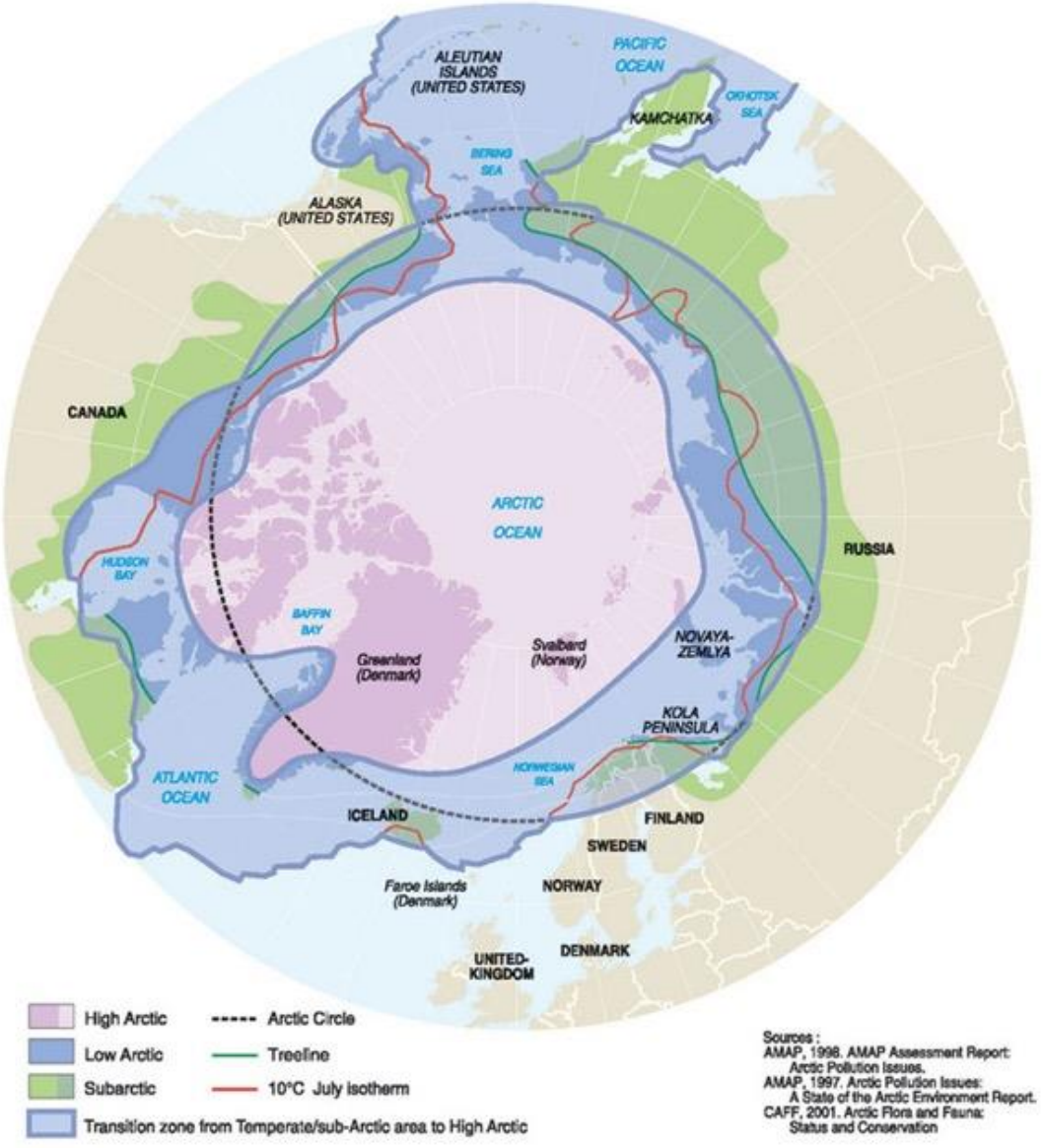

Figure 1.1 Definitions of Arctic boundaries (Arctic Centre, University of Lapland, n.d.).

1.3 Algae in Arctic Ecosystems

Algal communities in Arctic marine ecosystems can be roughly divided into three categories: ice algae, pelagic phytoplankton, and benthic algae. Each of these groups are photosynthetic organisms with the main differences being that ice algae thrive attached to a substrate, phytoplankton exist independently in the water column and benthic algae grow in or on marine sediment (Andersen and Lewin, 1999; CAFF, 2021; Encyclopaedia Britannica, 2019; Gradinger, 2020; Lovejoy, 2020). Both groups are principally composed of diatoms during the spring bloom (Campbell et al., 2015; Rózańska et al., 2009; Von Quillfeldt et al., 2003), sometimes making up greater than 90% of all species in ice-algal communities (Michel et al., 2002). Algae play a vital role in Arctic ecosystem structure and function through a variety of mechanisms (Arrigo, 2014). Biochemically, their roles in the food web, primary production, and carbon cycling identify them as essential components of their biome. As diatoms represent such a significant portion of these species and, as discussed below, contribute significantly to ecosystem function, these are the focus of this thesis.

1.3.1 Biogeochemical Roles

Marine food webs are driven by photoautotrophs (which include algae) that convert solar energy into organic compounds (Chassot et al., 2010). This class of organisms is responsible for over 90% of net primary productivity (NPP) in the global ocean (Cebrian and Duarte, 1996). In turn, these organic compounds provide a food source for larger, more complex organisms (Chassot et al., 2010). Cold temperatures and seasonally insufficient light conditions lead to slower growth and limit the rate of primary production in the Arctic relative to temperate and tropical ecosystems. However, within this slow-growth regime, algae play a significant ecological role for other organisms in the food web. Sea-ice algae in particular can be responsible for up to 57% of all NPP in the central Arctic Ocean under multi-year ice (Fernández-Méndez et al., 2015; Gosselin et al., 1997). Biophysical modelling programs have produced similar data, estimating the contribution of ice algae to equate to 40% locally and 7.5% for the Arctic as a whole (Dupont, 2012). The fat-rich pulse of production during a time of

limited food availability, providing sustenance for copepods and other zooplankton, is an essential service they provide to this region (Søreide et al., 2010).

Bradstreet and Cross (1982) were among the earliest to quantify the importance of specific trophic relationships within and surrounding Arctic ice-associated ecosystems. Their research outlined the interdependency between sea-ice diatoms and sympagic zooplankton (mainly copepods and amphipods) who feed almost exclusively on these diatoms for parts of the year. Some species of zooplankton have adapted their life cycles and reproductive strategies to coincide with annual algal blooms (Daase et al., 2013). The timing of the annual ice algal bloom in mid-May is critical to the survival of these meio- and macrofauna who rely on this early food source for sustenance before the annual phytoplankton bloom in the summer months (Michel et al., 1996; Søreide et al., 2010). As ice melts over the course of the spring and summer, algal aggregates descend into the water column and settle on the seabed where they are consumed by benthic invertebrates (Boetius et al., 2013).

The Canadian Arctic is home to almost 20% of all marine mammal species on Earth (Darnis et al., 2012). The Committee on the Status of Endangered Wildlife in Canada (COSEWIC) has established many of these mammals as endangered, threatened, or of concern, including populations of *Delphinapterus leucas* (beluga whale), *Odobenus rosmarus rosmarus* (Atlantic walrus), *Balaena mysticetus* (bowhead whale), *Orcinus orca* (killer whale), *Monodon monoceros* (narwhal), and *Ursus maritimus* (polar bear) (Government of Canada, 2018). As shown in Figure 1.2 (Darnis et al., 2012) the entire Arctic food web including these at-risk species relies heavily upon the services offered by ice algae and phytoplankton through these trophic interactions. Changes in algal community structure or abundance would have detrimental ramifications up the food web, ultimately having negative implications for human settlements in the North that rely on these species for their livelihood. For example, Gosselin et al. (1997) estimated the production rate of dissolved organic carbon (DOC) in the Arctic Ocean to be 36 mg C m^{-2} day^{-1} (approximately 15 g C m^{-2} annually). Specifically, pennate diatoms have accounted for up to 77% of algal species in landfast and pack ice and are a major contributor of DOC to their ecosystems (Poulin et al., 2011). Through direct and indirect feeding strategies, diatoms are considered to

be the main source of carbon for the ecologically keystone species *Boreogadus saida* (Arctic cod) (Bradstreet and Cross, 1982; Kohlbach et al., 2017), which are recognized to be some of the most abundant fish in this environment (David et al., 2016; Gradinger and Bluhm, 2004). Arctic cod make up the vast majority of marine mammal and bird diets, including *Uria lomvia* (thick-billed murre), *Fulmarus glacialis* (fulmar), *Phoca hispida* (ringed seal), and *Monodon monoceros* (narwhal) (Bradstreet and Cross, 1982).

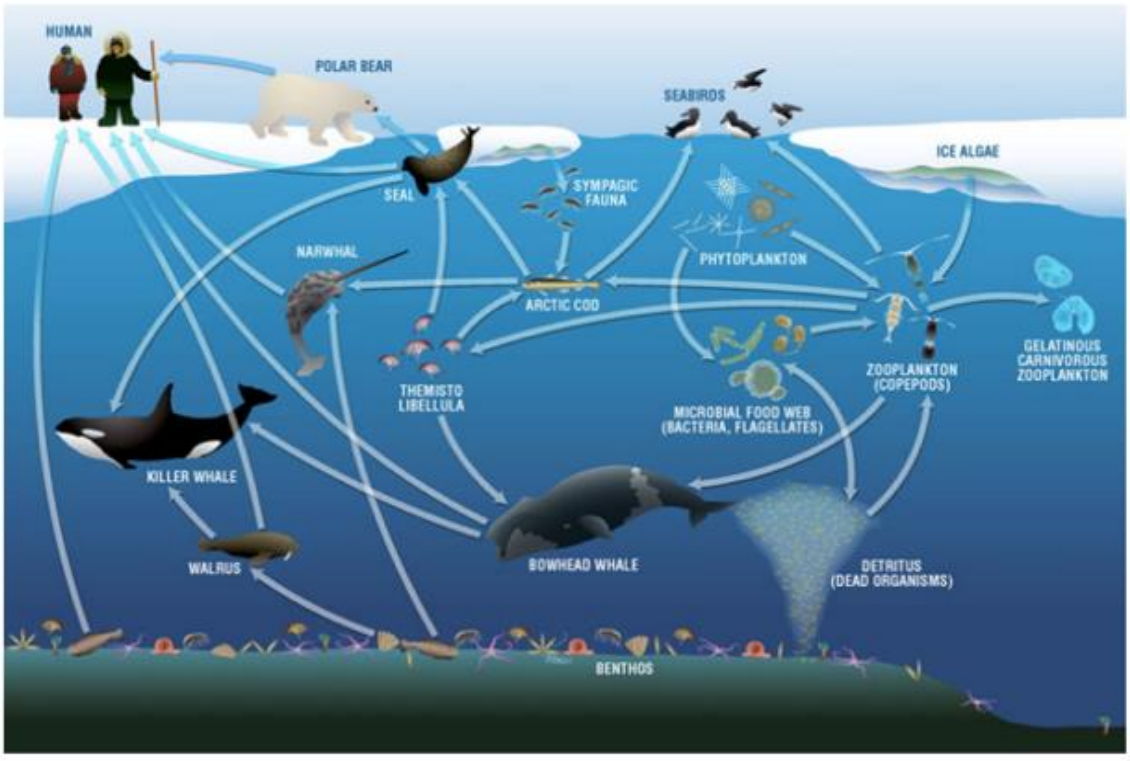

Figure 1.2 Canadian Arctic marine food web demonstrating the importance of ice algae and phytoplankton as drivers of the ecosystem (taken from Darnis et al., 2012).

1.3.2 Ice Algae

Sea ice covers more than 6% of all oceanic surface area on Earth at its annual maximum, with Arctic sea ice specifically extending to over 14 million square kilometres in winter in recent years (Arrigo, 2014; Scott and Hansen, 2016). The enormity of this habitat provides a vast range over which sympagic species can thrive. Ice algae are the drivers of life at a critical time of year within these ecosystems as described in Section 1.3.1; specifically, the communities of diatoms that are abundant in the bottom layer of seasonally-formed first-year ice (Ambrose et al., 2005). They dominate the ice-water

interface of this skeletal layer and are also found in interstitial brine channels (Figure 1.3, Arrigo, 2014; Poulin et al., 2014) that host a variety of micro- and meiofauna (Nozais et al., 2001).

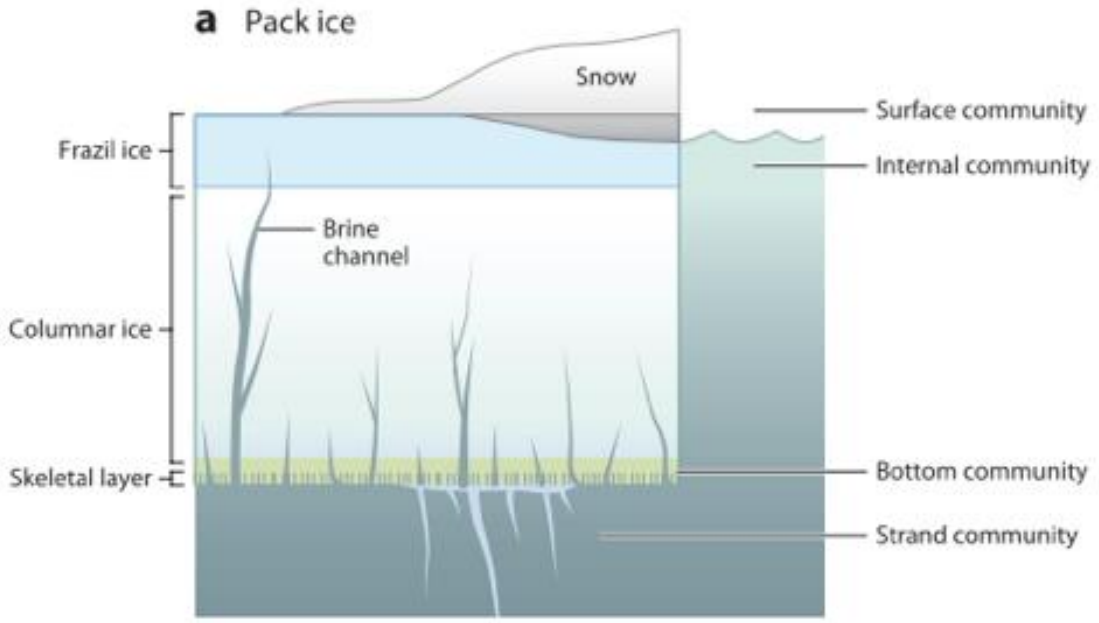

Figure 1.3 Depiction of skeletal layer and brine channels within Arctic pack ice where ice algae form communities (taken from Arrigo, 2014).

Before the seasonal melt, zooplankton feed on the algae attached to the skeletal layer of first-year ice and during the melt pelagic grazers are present to consume the cells as they descend into the water column (Michel et al., 1996). Not only is timing availability an important factor in ice algae as a food source, but these microalgae are producers of long chain omega-3 fatty acids which are required for survival by the vast majority of marine species (McMahon et al., 2006; Søreide et al., 2010). These fatty acids are also a dietary staple for the populations of the keystone bivalve *Mytilus trossulus* at their northern limit in Greenland (Thyrring et al., 2017). Further research has evidenced sea-ice derived carbon directly in polar bear tissue, highlighting the importance of trophic level-transferred carbon to the ecosystem as a whole (Brown et al., 2018). This demonstrates that their role in the food web is likely irreplaceable by

other organisms and that a decrease in their abundance would have widespread ramifications for the system.

Some laboratory research, though limited in volume and scope, has been conducted using the Arctic ice algae *Nitzschia frigida*. These data have informed our knowledge of physiological responses of this diatom under varying environmental conditions, providing insight as to how the changing Arctic climate may influence its abundance in sea ice. Aumack and Juhl (2015) assessed settling rates of *N. frigida* under limited nutrient and light conditions and found that cultures incubated under low light with limited silicate had significantly different settling characteristics than controls. Further to the discussion in Section 1.3.1 regarding the importance of the seasonal timing of the ice-algal bloom, the settling rate of these algae is key to the overall phenology of the Arctic ecosystem. The impact of extreme temperatures on growth and productivity has also been assessed with a temperature of -8ºC being established as the lower limit for growth (Aletsee and Jahnke, 1992). Studies of the physiological mechanisms by which specific systems within the cell function relative to factors in their environment have been investigated. Mitchell and Beardall (1996) found that *N. frigida* has a carbon-concentrating mechanism that may provide some unknown ecological advantage. Other research has highlighted the potential sensitivity of this diatom to stressors related to climate change such as ocean acidification and varying light conditions (Kvernvik et al., 2020).

1.3.3 Socioeconomic Importance

As previously discussed, changes to the structure and abundance of algal communities in the Arctic food web would negatively impact higher-trophic level mammals. Charismatic megafauna are dependent on algal primary production energy transfer and are, in turn, depended upon by northern communities as part of their social, cultural, economic and dietary routines (Nuttall et al., 2005). Indigenous peoples rely on marine mammals and fish for sustenance as the importation of other food sources is often difficult, unattainable, or unaffordable (Hossain et al., 2018; Nuttall et al., 2005). Arctic cod, which have been identified as a keystone species that rely on the services of algae, have historically been important in northern commercial fisheries throughout

Canada, Greenland, Iceland, and Norway and continue to be fished for subsistence by Indigenous communities (Kohlbach et al., 2017; Nuttall et al., 2005).

1.4 Threats to Arctic Ecosystems

Arctic ecosystems are becoming increasingly threatened over time due to increases in population, shipping and other marine traffic, and resource development and expansion. Statistics Canada census data demonstrated a 12.7% population increase in Nunavut between 2011 and 2016, achieving the highest rate of growth among all provinces and territories (Statistics Canada, 2016). The Arctic Council reported a 25% increase in the number of ships entering the Arctic during a six-year period between 2013 and 2019 (PAME, 2020). Total dollar estimates for all Arctic resources including fisheries, minerals, and oil are estimated at over $290 billion per year (O'Garra, 2017). These activity increases have vast implications for Arctic ecosystems as a whole, but especially for aquatic primary producers.

1.4.1 Wastewater Effluent

1.4.1.1 Wastewater treatment in northern regions

Wastewater management in northern regions faces many challenges. The remote location of communities in the Canadian North is an obstacle to the transportation of construction materials for building infrastructure, and freeze-thaw cycles combined with a harsh climate pose a challenge for continuous operation of machinery required for mechanical systems (Krumhansl et al., 2015). Thus, the majority of wastewater treatment in Arctic communities is accomplished using passive methods or by direct discharge into the sea with no treatment (Gunnarsdóttir et al., 2013; Hermanson et al., 2010; Johnson, 2008; Krumhansl et al., 2015).

Passive wastewater treatment systems are often in the form of waste stabilization ponds, natural or man-made lagoons, wetland systems, settling tanks (if the infrastructure exists), or a combination of these (Chaves-Barquero et al., 2016; Gunnarsdóttir et al., 2013; Krumhansl et al., 2015; Sonune and Ghate, 2004). These primary treatment mechanisms are designed to remove organic and inorganic materials

via physical processes. Waste is allowed to settle so that heavier materials sink and lighter materials such as oils and grease float to the surface. Natural microbial degradation (secondary treatment) sometimes follows before effluent is discharged into the environment. Although these practices are rudimentary, it significantly improves the water quality of primary effluent (i.e., reduction of up to 50% of biochemical oxygen demand [BOD], removal of up to 70% of total suspended solids, and removal up to 65% of oils and grease) (Sonune and Ghate, 2004). An example of typical lagoon-treated sewage infrastructure exists in Fort Resolution, NWT. More than 45,000 litres of raw sewage are collected from the municipality weekly and trucked to a sewage lagoon where aerobic and anaerobic biodegradation act as the first step in sewage treatment. From there, sewage infiltrates the soil before reaching the groundwater table, where harmful bacteria and excess nutrients are biodegraded. After a number of years the treated sewage enters a wetland system and is subsequently discharged into Great Slave Lake, though in other regions the final discharge point could be marine or freshwater (Johnson, 2011). Although this is an improvement from untreated sewage being directly discharged into aquatic environments, this system is still insufficient. For example, in 2007/2008, more than 60% of water samples taken from municipal wastewater effluent in northern communities did not satisfy the national performance standards set out by the Canadian Council of Ministers of the Environment (CCME) via the Canada-Wide Strategy for the Management of Municipal Wastewater Effluent (CCME, 2009; Johnson, 2011). These minimum requirements dictate upper limits in effluent of 25 mg/L carbonaceous BOD, 25 mg/L total suspended solids, and 0.02 mg/L total residual chlorine (CCME, 2009).

As of 2008, only 3 of 53 Inuit communities in Nunavut used mechanical systems, all of which have faced operational challenges often leading to temporary shutdowns. These mechanized operations are mainly to aid in preliminary or primary treatment that screens out large solids and particulate matter, except for the more advanced rotating biological contactor in Pangnirtung (Johnson, 2008).

1.4.1.2 Contaminants in Arctic wastewater

The risks of wastewater contaminants to ice algae and other members of ice-associated communities may be greater than elsewhere in the ecosystem due to their choice of habitat. Wind-driven and oceanic currents can be dampened when the ice layer acts as a boundary between currents and the organisms in under-ice or bottom-ice communities, so contaminant exposure can be much longer than in areas being constantly mixed and diluted by natural processes (Chapman and Riddle, 2003). It has also been reported that the degradation time for pharmaceuticals in Arctic environments is much slower than in warmer ecosystems, leading to longer exposure times for organisms (Kallenborn et al., 2008).

Kallenborn et al. (2018) published a review on pharmaceuticals and personal care products (PPCPs) found in Arctic ecosystems wherein domestic/municipal and sewage wastes were the primary sources. In marine environments, 11 types of non-steroidal anti-inflammatory and antipyretic analgesic drugs (NSAIDs), ten types of antidepressants/selective serotonin re-uptake inhibitors (SSRIs), four types of antimicrobials, 14 types of antibiotics, and two types of stimulants were found. Specifically, the highest levels of the anticonvulsant drug carbamazepine were measured in Arctic regions of Canada (47.7 ng/L in Hudson Bay). Environmentally-relevant mixtures of pharmaceuticals have demonstrated toxicity to algae specifically at concentrations in the low µg/L range. For example, Watanabe et al. (2015) reported a clarithromycin 72-hour laboratory IC50 for *Pseudokirchneriella supcapitata* of 6.9 µg/L.

Little research has been done in Canada to quantify emerging contaminant presence and levels in wastewater from Arctic settlements, though there has been increased interest in recent years (Chaves-Barquero et al., 2016; Gunnarsdóttir et al., 2013; Hayward and Jamieson, 2015; Krumhansl et al., 2015; McCarter et al., 2017; Stroski et al., 2020). Limitations in sample collection of wastewater effluent exist in the Arctic for the same reasons as the described barriers for implementing wastewater treatment infrastructure: seasonal inaccessibility, remote locations, and cost of time and resources all present obstacles to completing these exercises. Chaves-Barquero et al. (2016) detected atenolol (97 ng/L), carbamazepine (428 ng/L), clarithromycin (9 ng/L),

metoprolol (2 ng/L), sulfamethoxazole (274 ng/L) and trimethoprim (26 ng/L) in effluent samples from a lagoon-wetland system in Cambridge Bay, Nunavut, though it was determined in this study that the contaminants were not present in high enough concentrations to present a risk to marine life based on current toxicity data for temperate species as the only data available. However, it was noted that the majority of toxicity data used in calculating hazard quotients were from freshwater organisms as there are few marine organism data points. This highlights the need for conducting standard toxicity tests with representative marine species in order to fill this knowledge gap. This study also measured phosphorus and nitrogen concentrations in various locations relative to the wastewater treatment facility and found that the lagoon system was not successful in removing excess nutrients from the effluent (Chaves-Barquero et al., 2016), leading to significant impacts on local primary production (Back et al., 2021).

Stroski et al. (2020) detected numerous pharmaceuticals in Arctic wastewater, including atenolol, carbamazepine, metoprolol, and naproxen; however, the majority of samples taken from receiving waters showed concentrations that did not pose a risk to aquatic life. Concentrations of per- and polyfluorinated alkyl substances (PFAS) were also studied and levels of these compounds were found to be locally significant in Cambridge Bay receiving waters during wastewater release, but below levels of detection after release. Other studies have been done in non-Canadian Arctic regions (Gunnarsdóttir et al., 2013; Leclerc et al., 2002; Weigel et al., 2004; that have found concentrations of caffeine, ibuprofen, *N,N*-diethyl-3-toulamide (DEET, the active ingredient in most insect repellents), pathogens, and parasites in wastewater effluent.

Persistent organic pollutants (POPs) have been found in domestic sewage from Inuit communities in northern Canada (Hermanson et al., 2010). These contaminants, particularly polychlorinated biphenyls (PCBs) and perfluorinated compounds (PFCs) bioaccumulate in marine, aquatic, and terrestrial animals that Inuit populations depend on as a traditional food source (Ostertag et al., 2009). Aside from the bioaccumulation potential in larger mammals, PCBs have also demonstrated direct toxicity to temperate algal species at concentrations as low as 0.34 µM (Halm-Lemeille et al., 2014).

1.4.2 Resource Development

Human activities in this region such as mineral extraction, hydrocarbon exploration, and shipping (AMAP, 2017; Aulanier et al., 2017; Hsiao, 1978) all pose potential threats to flora and fauna through the release of contaminants into aquatic ecosystems. In 2008, the United States Geological Survey estimated that up to 90 billion barrels of oil (~14 trillion litres), 1,669 trillion cubic feet (~47 trillion m^3) of natural gas, and 44 billion barrels (~7 trillion litres) of natural gas liquids are likely to remain undiscovered in the Arctic (United States Department of the Interior, 2008). Their estimate indicated that 84% of these resources are located offshore, potentially posing a threat to marine ecosystems upon their imminent exploitation if spills or leakage were to occur (United States Department of the Interior, 2008). These processes bring to light several contaminants of concern; namely polycyclic aromatic hydrocarbon compounds (PAHs) that are present in crude oil which can become more bioavailable to organisms when oil spill response chemicals are applied for remediation and response (Schein et al., 2009). Dispersed oil has demonstrated growth inhibition and decreased motility in the algae *Isochrysis galbana* specifically (Garr et al., 2014). Components of crude oil alone can cause acute mortality, behavioural changes, and developmental and reproductive defects in all trophic levels of marine organisms (Bejarano et al., 2017; Hansen et al., 2011; Hsiao, 1978; Nørregaard et al., 2014, Olsen et al., 2011; Philibert et al., 2016). Acute and chronic toxicity of total petroleum hydrocarbons (TPHs) at environmentally-relevant concentrations (i.e., as low as 1.6 mg/L) has been observed in keystone Arctic species including the copepod *Calanus glacialis* and the ray-finned fishes *Boreogadus saida* and *Myoxocephalus sp.* (Gardiner et al., 2013; Hansen et al., 2014), all of which rely on algae for survival. Direct toxicity to standard algal test species (i.e., *Pseudokirchneriella subcapitata*) has also occurred from PAH exposure at concentrations in the µg/L range (Bragin et al., 2016).

1.4.3 Climate Change

Changes in climate including warming ocean temperatures, reduced sea ice cover, and increased light intensity in surface waters, plus the amplification of these issues via positive feedback loops such as the ice and snow melt/surface albedo coupling, have already had direct, measurable impacts on Arctic ecosystems. Quantified responses

include community and regime shifts, changes in behaviour and phenology, growth, condition, and abundance of various species, and shifts in geographic range (Wassmann et al., 2010). These climatic changes have the potential to exacerbate organismal responses to contaminant exposure by creating a multiple-stressor environment. The steady annual reduction in Arctic sea ice cover has dramatic implications for the future of sea ice algae, specifically, introducing even further uncertainty in extrapolating how upper-trophic level biota may respond to these events (Comiso et al., 2008).

1.5 Critical Review

It has been noted in the literature that there is little comprehension of Arctic species' sensitivity to contaminants and that this lack of information may result in harm to this ecosystem (Chapman and Riddle, 2003). Though there has been a growing focus to quantify the impact of anthropogenic activities on Arctic ecosystems in recent years, most regulatory guidelines and risk assessments are built primarily around temperate species. To help address this concern, a critical review to assess reported effects for these species in relation to current regulatory guidelines, quantify knowledge and methodological gaps, and identify future research needs for laboratory testing was performed. To accomplish this, an objective evaluation of the literature was employed to quantify the strengths and relevance of published studies to determine the utility of the data. A uniform set of criteria to score each study was developed, allowing an objective comparison across experiments. Data from all publicly available toxicity tests performed on aquatic Arctic algae (n = 22 distinct experiments), invertebrates (n = 167), vertebrates (n = 149) and microbial consortiums (n = 1) were included.

1.6 Algal Toxicity Tests

To help address these threats, ERAs are employed. Environmental risk is characterized through the comparison of contaminant effect data and environmentally-relevant concentrations of the contaminants to be assessed. Risk assessors use ecotoxicology data to aid in regulatory decision making for environmental protection for substances including pesticides, industrial chemicals, and wastewater effluent, including to define threshold concentrations of contaminants that are allowed to be present in aquatic bodies (Government of Canada, 2019). Algae are often the most sensitive organism type to contaminant exposure, largely due to their size, which renders them excellent representatives for defining thresholds for ecosystem protection to ensure all other organism classes will be protected. Still, maintaining cultures of Arctic organisms in a laboratory for testing is often more difficult than maintaining temperate ones due to specialized adaptations to their environment. For example, many standard species used in toxicology research are from temperate regions and can be reared at near-room

temperature (i.e., *Daphnia*, *Skeletonema costatum,* and *Oncorhynchus mykiss*) while Arctic species require chilled water and specialized photoperiods. These reasons may contribute to the lack of published protocols designed specifically for Arctic toxicity tests. Though a few Arctic species have been tested through the modification of standard protocols (i.e., *Porosira glacialis* in Camus et al., 2015), there are no standard methods that are widely used. To date, only 44 ecotoxicology studies with Arctic species have publicly available data that can be used in ERA. There is a need to incorporate the use of Arctic organisms in standard toxicity tests due to knowledge gaps in their sensitivity to contaminants that are released into their environment.

1.6.1 Laboratory Assay Features, Strengths, And Weaknesses

Toxicological values including no observed effect concentrations (NOECs), lowest observed effect concentrations (LOECs), concentrations where effects are observed in x% of a population (ECx) and more specifically, lethal or inhibitory effects (LCx or ICx) are derived from laboratory studies where organisms are exposed to the contaminant of concern. In general, laboratory studies are designed to be conservative in their results and are a first-tier screening tool for observing direct toxicity. Exposure conditions are not meant to mimic scenarios that take place in the environment, but instead to provide a worst-case scenario to ensure that similar concentrations in the environment will always be protective of the ecosystem as a whole. The sole question being asked by these assays is typically whether the contaminant is toxic to the organism being tested without considering how outside agents may act on these processes in nature. For example, lab assays typically use filtered seawater to avoid fouling over the duration of the exposure and limit opportunities for other variables to be introduced that could render it difficult to discern effects due to the contaminant of concern alone. In nature, there are undoubtedly bacteria and other organisms present in the exposure scenario that could reduce toxicity by breaking down the compound, but the presence of these in a laboratory assay interferes with the ability to answer the question of toxicity to the test organism. It has also been suggested that maintaining laboratory cultures of microalgae over long periods of time may result in phenotypic and genotypic changes as they adapt to their environmental conditions (Lakeman et al.,

2009). This process of evolutionary adaptation could influence the sensitivity and thus experimental results of a laboratory culture relative to field populations of the same species. Though these issues can be considered as weaknesses of the laboratory test from the perspective of environmental relevance, it is a necessary tool to avoid spending time and resources investigating effects in an environmentally realistic exposure scenario if the compound is not toxic in a worst-case exposure.

An important strength of the laboratory assay that is difficult to accomplish through field exposures is the ability to replicate scenarios across space and time to increase confidence in observed effects. The use of standard methods, test species, and reference toxicants allows labs to determine the replicability and reproducibility of their results. Environment and Climate Change Canada (ECCC) suggests that aquatic toxicity testing laboratories conduct reference toxicant testing at least once per month. With these data, comparisons of relative sensitivities between different test species can be made to further determine suitability for future testing.

Algal assays have been critiqued for employing culture conditions that are not representative of environmental conditions. For example, continuous light for a 72- to 96-hour duration does not occur in nature where these species exist; however, it is required to ensure that sufficient growth is achieved within a short time frame to discern effects between treatments at the end. Similarly, water temperature in natural environments inhabited by these temperate species could be consistently much cooler than the 20˚C range, but colder temperatures may slow down metabolic processes, lengthening the duration required to observe effects and increasing overall effort. The strengths of laboratory toxicity tests are as a screening tool to determine whether effects of specific contaminant exposure warrant further investigation, not to be exact proxies for environmental exposure.

1.6.2 Current Methods For Algal Toxicity Testing

Algal toxicity testing protocols are standardized to ensure that tests are run the same way across spatial and temporal bounds. This uniformity allows data to be transferrable between laboratories and over time. Similar experimental design and unvarying environmental parameters typically provide consistent results, giving

researchers more confidence in the data they produce. This also limits the frequency that tests with the same chemicals or test organisms need to be performed, reserving resources for other projects.

Standard methods provide guidelines for all aspects of the test including acceptable ranges for environmental parameters, experimental design, test vessels and organisms, test duration, validity criteria, and statistical analyses required to obtain values for specified endpoints (Table 1.1). Widely accepted methods for algal tests have been put forth by many agencies, but the most widely accepted and commonly applied (the American Society for Testing and Materials [ASTM International], Environment and Climate Change Canada [ECCC], the National Institute for Water and Atmospheric Research [NIWA], the Organization for Economic Cooperation and Development [OECD], and the United States Environmental Protection Agency [US EPA]) were drawn from for this study. Each have their own advantages and drawbacks which have been formally compared and critiqued in the literature (Eisentraeger et al., 2003; Fitzgerald, 1975; Weiss, 1976;). The major differences between methods are the volume requirements for experimental units (EUs) as determined by test vessel size. The microplate technique has been praised for its low volume requirements of both test solution and algal inoculum, which ultimately leads to a reduced incubation space requirement for the test as a whole (ECCC, 2007; NIWA, 1998). The bottle test, or “flask method”, named for the use of Erlenmeyer flasks as test vessels, is employed elsewhere and requires significantly more space and resources to complete than the microplate technique. Other benefits of the microplate include the disposable nature of test vessels and pipette tips, limiting the risk of contamination from insufficiently cleaned glassware and reducing the time required for rinsing and sterilizing experimental vessels (NIWA, 1998). Despite these benefits, the use of a microplate limits the substances that can be tested for two reasons: volatile compounds may affect the algae in adjacent wells, and the polystyrene composition of the vessel may interact with test material as opposed to inert glass. Both of these disadvantages are addressed through the use of the flask method.

Table 1.1 Summary of selected standard methods for algal toxicity testing.

	NIWA (1998)	OECD (2011)	ASTM (2012)	US EPA (2012)	ECCC (2007)
Species	*Dunaliella tertiolecta*	*Pseudokirchneriella subcapitata* *Desmodesmus subspicatus* *Aphanizomenon flos-aquae* *Synechoccus leopoliensis* *Navicula pelliculosa*	Various	*P. subcapitata* *Skeletonema costatum* *N. pelliculosa*	*P. subcapitata*
Temp. (°C)	20 ± 2	21 to 24 ± 2*	20 to 24 ± 2*	20 to 24 ± 2*	24 ± 2
Salinity (ppt)	35	NA	24-35	30 ± 5	NA
Light (µmol photo m^{-2} s^{-1})	200	40-200*	30-90	60	50-62*
Inoculum (cells/mL)	10,000	2,000 - 500,000*	10,000-50,000*	10,000	10,000
Endpoint	IC50 biomass	ECx biomass	IC50 biomass	IC50 biomass	IC50 cell yield

*Species dependent

There are several standard test species that have been widely accepted as representatives for both freshwater and marine environments. As evidenced in Table 1.1, all acceptable species inhabit temperate regions with culturing temperatures in the range of 20˚C. A range of culturing methods are described according to test species, as well as various inoculum concentrations, growth criteria, and acceptable endpoints.

Reference toxicants commonly used in algal tests are comprised of a variety of compound groups from phenols to heavy metals and salts (ECCC, 1990). ECCC has outlined specific criteria that should be met by a chemical to be suitable for reference toxicity testing based on eight criteria including solubility, stability, intralaboratory water quality effects, and ease of chemical analysis. The 96-hour *Selenastrum capricornutum* growth inhibition test uses phenol, hexavalent chromium, copper or zinc for reference toxicants.

1.7 Objectives

The broad purpose of this study was to further the understanding of Arctic species' sensitivity to contaminants in order to inform ERA in the region. This was accomplished through the following objectives:

1 Conduct a critical review on the current state of Arctic ecotoxicology.
- Compile existing data from published studies to form a database.
- Conduct a formal critique of this literature to assess the reliability of the data for use in ERA.
- Identify knowledge gaps that exist in the field to outline needs for further research.

2 Optimize culture conditions for the toxicity testing of Arctic algal species.
- Identify representative marine species of Arctic algae that are amenable to routine laboratory culturing.
- Design and conduct an experiment to optimize culture conditions such that toxicity tests can be performed efficiently.

3 Develop a protocol for toxicity testing with candidate species.

- Determine the repeatability of toxicity tests following modified standard protocols with these relevant species at the Huntsman Marine Science Centre (HMSC) in St. Andrews, New Brunswick.
- Determine replicability by providing a protocol for other laboratories to use where results can eventually be compared across institutions.
- Provide a statistically sound dataset of toxicological responses that can be used in future ERAs for the Arctic region ensuring they are accurate and robust.

1.8 References

Aletsee, L. and Jahnke, J. (1992). Growth and productivity of the psychrophilic marine diatoms *Thalassiosira antarctica* Comber and *Nitzschia frigida* Grunow in batch cultures at temperatures below the freezing point of sea water. *Polar Biology, 11*, 643-647.

Ambrose, W. G., Von Quillfeldt, C., Clough, L. M., Tilney, P. V. R., and Tucker, T. (2005). The sub-ice algal community in the Chukchi sea: Large- and small-scale patterns of abundance based on images from a remotely operated vehicle. *Polar Biology*, *28*(10), 784–795.

American Society for the Testing of Materials. (2012). Standard Guide for Conducting Static Toxicity Tests with Microalgae.

Andersen, R. A. and Lewin, R. A. (1999). *Algae*. In *Encyclopaedia Britannica Online.* Retrieved from http://www.britannica.com/science/algae.

Arctic Centre, University of Lapland. (n.d.). Definitions of the Arctic region. Retrieved 06 February 2019 from https://www.arcticcentre.org/EN/communications/arcticregion/Maps/definitions#.

Arctic Council. (2015). Member states. Retrieved 05 February 2019 from https://arctic-council.org/index.php/en/about-us/member-states.

Arctic Council. (2020). About. Retrieved 22 July 2020 from https://arctic-council.org/en/about/.

Arctic Monitoring and Assessment Programme (AMAP). (2017). *Adaptation Actions for a Changing Arctic: Perspectives from the Bering-Chukchi-Beaufort Region*. *Arctic Monitoring and Assessment Programme (AMAP)*.

Arrigo, K. R. (2014). Sea Ice Ecosystems. *Annual Review of Marine Science*, *6*(1), 439–467.

Aulanier, F., Simard, Y., Roy, N., Gervaise, C., and Bandet, M. (2017). Effects of shipping on marine acoustic habitats in Canadian Arctic estimated via probabilistic modeling and mapping. *Marine Pollution Bulletin*, *125*(1–2), 115–131.

Aumack, C. F. and Juhl, A. R. (2015). Light and nutrient effects on the settling characteristics of the sea ice diatom *Nitzschia frigida. Limnology and Oceanography, 60*(3), 765-776.

Back, D.-Y., Ha, S.-Y., Else, B., Hanson, M., Jones, S. F., Shin, K.-H., Tatarek, A., Wiktor, J. M., Cicek, N., Alam, S., Mundy, C. J. (2021). On the impact of wastewater effluent on phytoplankton in the Arctic coastal zone: A case study in the Kitikmeot Sea of the Canadian Arctic. *Science of the Total Environment, 764,* 143861.

Bejarano, A. C., Gardiner, W. W., Barron, M. G., Word, J. Q. (2017). Relative sensitivity of Arctic species to physically and chemically dispersed oil determined from three hydrocarbon measures of aquatic toxicity. *Marine Pollution Bulletin, 122*(1-2), 316-322.

Bellamy, J. (2010). Definitions of the Arctic region [Geologic map]. Retrieved from https://www.amap.no/documents/doc/definitions-of-the-arctic-region/248.

Boetius, A., Albrecht, S., Bakker, K., Bienhold, C., Felden, J., Fernández-Méndez, M., Hendricks, S., Katlein, C., Lalande, C., Krumpen, T., Nicolaus, M., Peeken, I., Rabe, B., Rogacheva, A., Rybakova, E., Somavilla, R., Wenzhöfer, F., and RV Polarstern ARK27-3-Shipboard Science Party. (2013). Export of algal biomass from the melting Arctic sea ice. *Science, 339*(6126), 1430-1432.

Bradstreet, M. S. W., and Cross, W. E. (1982). Trophic relationships at High Arctic ice edges. *Arctic*, *35*(1), 1–12.

Bragin, G. E., Parkerton, T. F., Redman, A. D., Letinksi, D. J., Butler, J. D., Paumen, M. L., Sutherland, C. A., Knarr, T. M., Comber, M., den Haan, K. (2016). Chronic toxicity of selected polycyclic aromatic hydrocarbons to algae and crustaceans using passive dosing. *Environmental Toxicology and Chemistry, 35*(12), 2948-2957.

Brown, T. A., Galicia, M. P., Thiemann, G. W., Belt, S. T., Yurkowski, J., and Dyck, M. G. (2018). High contributions of sea ice derived carbon in polar bear (Ursus maritimus) tissue. *PloS ONE*, *13*(1), e01916311–13.

Campbell, K., Mundy, C. J., Barber, D. G., and Gosselin, M. (2015). Characterizing the sea ice algae chlorophyll a-snow depth relationship over Arctic spring melt

using transmitted irradiance. *Journal of Marine Systems*, *147*, 76–84.

Camus, L., Brooks, S., Geraudie, P., Hjorth, M., Nahrgang, J., Olsen, G. H., Smit, M. G. D. (2015). Comparison of produced water toxicity to Arctic and temperate species. *Ecotoxicology and Environmental Safety, 113,* 248-258.

Canadian Council of Ministers of the Environment. (2009). Canada-wide Strategy for the Management of Municipal Wastewater Effluent.

Cebrian, J., and Duarte, C. M. (1996). The fate of marine autotrophic production. *Limnology and Oceanography*, *41*(8), 1758–1766.

Chapman, P. M., and Riddle, M. J. (2003). Missing and needed: Polar marine ecotoxicology. *Marine Pollution Bulletin*, *46*(8), 927–928.

Chapman, P. M., and Riddle, M. J. (2005). Polar marine toxicology – Future research needs. *Marine Pollution Bulletin*, *50*(9), 905–908.

Chassot, E., Bonhommeau, S., Dulvy, N. K., Mélin, F., Watson, R., Gascuel, D., and Le Pape, O. (2010). Global marine primary production constrains fisheries catches. *Ecology Letters*, *13*(4), 495–505.

Chaves-Barquero, L. G., Luong, K. H., Mundy, C. J., Knapp, C. W., Hanson, M. L., and Wong, C. S. (2016). The release of wastewater contaminants in the Arctic: A case study from Cambridge Bay, Nunavut, Canada. *Environmental Pollution*, *218*, 542–550.

Comiso, J. C., Parkinson, C. L., Gersten, R., Stock, L. (2008). Accelerated decline in the Arctic sea ice cover. *Geophysical Research Letters, 35*(1).

Conservation of Arctic Flora and Fauna (CAFF). (2021). *Algae from benthic samples.* Arctic biodiversity. Retrieved December 30, 2021 from https://www.arcticbiodiversity.is/index.php/finding/algae-from-benthic-samples.

Daase, M., Falk-Petersen, S., Varpe, Ø., Darnis, G., Søreide, J. E., Wold, A., Leu, E., Berge, J., Phillipe, B., Fortier, L. (2013). Timing of reproductive events in the marine copepod *Calanus glacialis*: a pan-Arctic perspective. *Canadian Journal of Fisheries and Aquatic Sciences*, *70*(6), 871–884.

Darnis, G., Robert, D., Pomerleau, C., Link, H., Archambault, P., Nelson, R. J., Geoffroy, M., Tremblay, J.-E., Lovejoy, C., Ferguson, S. H., Hunt, B. P. V., Fortier, L. (2012). Current state and trends in Canadian Arctic marine

ecosystems: II. Heterotrophic food web, pelagic-benthic coupling, and biodiversity. *Climatic Change, 115*(1), 179-205.

David, C., Lange, B., Krumpen, T., Schaafsma, F., van Franeker, J. A., and Flores, H. (2016). Under-ice distribution of polar cod Boreogadus saida in the central Arctic Ocean and their association with sea-ice habitat properties. *Polar Biology*, *39*(6), 981–994.

Dupont, F. (2012). Impact of sea-ice biology on overall primary production in a biophysical model of the pan-Arctic Ocean. *Journal of Geophysical Research: Oceans*, *117*(8), 1–18.

Eisentraeger, A., Dott, W., Klein, J., Hahn, S. (2003). Comparative studies on algal toxicity testing using fluorometric microplate and Erlenmeyer flask growth-inhibition assays. *Ecotoxicology and Environmental Safety 54*, 346-354.

Encyclopaedia Britannica, T. Editors of Encyclopaedia (2019, March 6). *Phytoplankton*. In *Encyclopaedia Britannica Online.* Retrieved from http://www.britannica.com/science/phytoplankton.

Environment and Climate Change Canada (ECCC). (1990). Guidance document on control of toxicity test precision using reference toxicants (Report EPS 1/RM/12).

Environment and Climate Change Canada (ECCC). (2007). Biological Test Method: Growth Inhibition Test Using a Freshwater Alga. EPS/1/RM/25 Second Edition.

Fernández-Méndez, M., Katlein, C., Rabe, B., Nicolaus, M., Peeken, I., Bakker, K., Flores, H., Boetius, A. (2015). Photosynthetic production in the central Arctic Ocean during the record sea-ice minimum in 2012. *Biogeosciences*, *12*(11), 3525–3549.

Fitzgerald, G. P. (1975). Factors affecting the algal assay procedure. US Environmental Protection Agency. OCLC15691319.

Gardiner, W. W., Word, J. Q., Word, J. D., Perkins, R. A., McFarlin, K. M., Hester, B. W., Word, L. S., Ray, C. M. (2013). The acute toxicity of chemically and physically dispersed crude oil to key arctic species under arctic conditions during the open water season. *Environmental Toxicology and Chemistry*, *32*(10), 2284–2300.

Garr, A. L. Laramore, S., Krebs, W. (2014). Toxic effects of oil and dispersant on marine microalgae. *Bulletin of Environmental Contamination and Toxicology, 93*,: 654-659.

Gosselin, M., Levasseur, M., and Wheeler, P. A. (1997). Deep Sea Research Part II: Topical Studies in Oceanography – New measurements of phytoplankton and ice algal production in the Arctic Ocean. *Deep Sea Research*, *44*(8).

Government of Canada. (2018). Species at Risk Public Registry. Retrieved 05 February 2019 from http://www.sararegistry.gc.ca.

Government of Canada. (2019). *Environmental Risk Assessment* [web page]. https://www.canada.ca/en/health-canada/services/consumer-product-safety/pesticides-pest-management/public/protecting-your-health-environment/pesticide-registration-process/reviews/environmental-risk-assessment.html.

Gradinger, R. (2020). Life in Arctic sea ice. In M. I. Goldstein and D. A. DellaSalla (Eds.), *Encyclopedia of the world's biomes.* Retrieved from https://www.sciencedirect.com/science/article/pii/B9780124095489124315.

Gradinger, R. R., and Bluhm, B. A. (2004). In-situ observations on the distribution and behavior of amphipods and Arctic cod (Boreogadus saida) under the sea ice of the High Arctic Canada Basin. *Polar Biology*, *27*(10), 595–603.

Gunnarsdóttir, R., Jenssen, P. D., Erland Jensen, P., Villumsen, A., and Kallenborn, R. (2013). A review of wastewater handling in the Arctic with special reference to pharmaceuticals and personal care products (PPCPs) and microbial pollution. *Ecological Engineering*, *50*, 76–85.

Halm-Lemeille, M.-P., Fard, E. A., Latire, T., Ferard, J.-F., Costil, K., Lebel, J.-M., Bureau, R., Serpentini, A. (2014). The effect of different polychlorinated biphenyl on two aquatic models, the green alga *Pseudokirchneriella subcapitata* and the haemocytes from the European abalone *Haliotis 37culate3737te. Chemosphere, 110*,: 120-128.

Hansen, B. H., Altin, D., Rørvik, S. F., Øverjordet, I. B., Olsen, A. J., and Nordtug, T. (2011). Comparative study on acute effects of water accommodated fractions of an artificially weathered crude oil on Calanus finmarchicus and Calanus

glacialis (Crustacea: Copepoda). *Science of the Total Environment*, *409*(4), 704–709.

Hansen, B., Altin, D., Bonaunet, K., Øverjordet, I. (2014). Acute toxicity of eight oil spill response chemicals to temperate, boreal, and Arctic species. *Journal of Toxicology and Environmental Health, Part A, 77*, 495-505.

Hayward, J. and Jamieson, R. (2015). Derivation of treatment rate constatnts for an Arctic tundra wetland receiving primary treated municipal wastewater. *Ecological Engineering, 82,* 165-174.

Hermanson, M. H., Jackson, D. S., and Johnson, G. W. (2010). History of Inuit community excretion of polychlorinated biphenyls recorded in sewage lake sediments. *Chemosphere*, *78*(11), 1322–1328.

Hossain, K., Raheem, D., and Cormier, S. (2018). Food Security Governance in the Arctic-Barents Region (pp. 15–32).: Springer International Publishing.

Hsiao, S. C. (1978). Effects of Crude Oils on the Growth of Arctic. *Environmental Pollution*, (17), 93–107.

Johnson, K. (2008). National Inuit Position Paper regarding the CCME Canada-Wide Strategy for the Management of Municipal Wastewater Effluent and Environment Canada's Proposed Regulatory Framework for Wastewater. Ottawa: Inuit Tapiriit Kanatami.

Johnson, K. (2011). Wastewater Treatment and Management in Northern Canada. Retrieved 08 February 2019 from https://issuu.com/cryofront/docs/northern_ww_treatment___management.

Kallenborn, R., Brorström-Lundén, E., Reiersen, L. O., Wilson, S. (2018). Pharmaceuticals and personal care products (PPCPs) in Arctic environments: Indicator contaminants for assessing local and remote anthropogenic sources in a pristine ecosystem in change. *Environmental Science and Pollution Research, 25*, 33001-33013.

Kallenborn, R., Fick, J., Lindberg, R., Moe, M., Nielsen, K. M., Tysklind, M., Vasskog, T. (2008). Pharmaceutical residues in Northern European environments: Consequences and perspectives. In K. Kummerer (Ed.), Pharmaceuticals in the Environment (pp. 61-74).

Kohlbach, D., Schaafsma, F. L., Graeve, M., Lebreton, B., Lange, B. A., David, C., Vortkamp, M., Flores, H. (2017). Strong linkage of polar cod (*Boreogadus saida*) to sea ice algae-produced carbon: Evidence from stomach content, fatty acid and stable isotope analyses. *Progress in Oceanography*, *152*, 62–74.

Krumhansl, K. A., Krkosek, W. H., Greenwood, M., Ragush, C., Schmidt, J., Grant, J., Barrell, J., Lu, L., Lam, B., Gagnon, G. A., and… Jamieson, R. C. (2015). Assessment of arctic community wastewater impacts on marine benthic invertebrates. *Environmental Science and Technology*, *49*(2), 760–766.

Kvernvik, A. C., Rokitta, S. D., Leu, E., Harms, L., Gabrielsen, T. M., Rost, B., Hoppe, C. J. M. (2020). Higher sensitivity towards light stress and ocean acidification in an Arctic sea-ice-associated diatom compared to a pelagic diatom. *New Phytologist, 226*(6), 1708-1724.

Lakeman, M. B. (2009). The strain concept in phytoplankton ecology. *Harmful Algae, 8,* 746-758.

Leclerc, H., Schwartzbrod, L., and Dei-Cas, E. (2002). Microbial agents associated with waterborne diseases. *Critical Reviews in Microbiology*, *28*(4), 371–409.

Lovejoy, C. (2020). Plankton of the open Arctic ocean. In M. I. Goldstein and D. A. DellaSalla (Eds.), *Encyclopedia of the world's biomes.* Retrieved from https://www.sciencedirect.com/science/article/pii/B9780124095489119797.

McCarter, C. P. R., Branfireun, B. A., and Price, J. S. (2017). Nutrient and mercury transport in a sub-arctic ladder fen peatland subjected to simulated wastewater discharges. *Science of the Total Environment*, *609*, 1349–1360.

McMahon, K. W., Ambrose, W. G. J., Johnson, B. J., Sun, M.-Y., Lopez, G. R., Clough, L. M., and Carroll, M. L. (2006). Benthic community response to ice algae and phytoplankton in Ny-Ålesund, Svalbard. *Marine Ecology Progress Series, 310*, 1–14.

Michel, C., Legendre, L., Ingram, R. G., Gosselin, M., and Levasseur, M. (1996). Carbon budget of sea-ice algae in spring: Evidence of a significant transfer to zooplankton grazers. *Journal of Geophysical Research: Oceans*, *101*(C8), 18345–18360.

Michel, C., Nielsen, T. G., Nozais, C., and Gosselin, M. (2002). Significance of

sedimentation and grazing by ice micro- and meiofauna for carbon cycling in annual sea ice (northern Baffin Bay). *Aquatic Microbial Ecology*, *30*(1), 57–68.

Mitchell, C. and Beardall, J. (1996). Inorganic carbon uptake by an Antarctic sea-ice diatom, *Nitzschia frigida. Polar Biology 16,* 95-99.

National Institute of Water and Atmospheric Research (NIWA). (1998). Appendix 1 (Marine algae [*Dunaliella tertiolecta*] chronic toxicity test protocol) to Standard Methods for Whole Effluent Toxicity Testing: Development and Application. Report no. MFE80205. NIWA Report for the Ministry for the Environment, Wellington, New Zealand.

Natural Resources Canada. (2017). The North. Retrieved 01 February 2019 from https://www.nrcan.gc.ca/earth-sciences/geography/atlas-canada/selected-thematic-maps/16886.

Nørregaard, R. D., Nielsen, T. G., Møller, E. F., Strand, J., Espersen, L., Møhl, M. (2014). Evaluating pyrene toxicity on Arctic key copepod species *Calanus hyperboreus. Ecotoxicology, 23*, 163-174.

Nozais, C., Gosselin, M., Michel, C., Tita G. (2001). Abundance, biomass, composition and grazing impact of sea-ice meiofauna in the North Water, northern Baffin Bay. *Marine Ecology Progress Series, 217*, 235-250.

Nuttall, M., Berkes, F., Forbes, B., Kofinas, G., Vlassova, T., and Wenzel, G. (2005). *Hunting, Herding, Fishing, and Gathering: Indigenous Peoples and Renewable Resource Use in the Arctic.* Arctic Climate Impact Assessment, 649-690.

O'Garra, T. (2017). Economic value of ecosystem services, minerals, and oil in a melting Arctic: A preliminary assessment. *Ecosystem Services, 24,:* 180-186.

Olsen, G. H., Smit, M. G. D., Carroll, J., Jaeger, I., Smith, T., Camus, L. (2011). Arctic versus temperate comparison of risk assessment metrics for 2-methyl-naphthalene. *Marine Environmental Research, 72,* 179-187.

Olsen, L. M., Laney, S. R., Duarte, P., Kauko, H. M., Fernández-Méndez, M., Mundy, C. J., Rösel, A., Meyer, A., Itkin, P., Cohen, L., Peeken, I., Tatarek, A., Różańska-Pluta, M., Wiktor, J., Taskjelle, T., Pavlov, A. K., Hudson, S. R., Granskog, M. A., Hop, H., … Assmy, P. (2017). The seeding of ice algal blooms in Arctic pack ice:

The multiyear ice seed repository hypothesis. *Journal of Geophysical Research: Biogeosciences*, *122*(7), 1529–1548.

Organization for Economic Cooperation and Development. (2011). OECD Guidelines for the Testing of Chemicals.

Ostertag, S. K., Tague, B. A., Humphries, M. M., Tittlemier, S. A., and Man, H. (2009). Estimated dietary exposure to fluorinated compounds from traditional foods among Inuit in Nunavut, Canada. *Chemosphere*, *75*(9), 1165–1172.

Philibert, D. A., Philibert, C. P. Lewis, C., Tierney, K. B. (2016). Comparison of diluted bitumen (dilbit) and conventional crude oil toxicity to developing zebrafish. *Environmental Science and Technology, 50*(11), 6091-6098.

Poulin, M., Daugbjerg, N., Gradinger, R., Ilyash, L., Ratkova, T., and von Quillfeldt, C. (2011). The pan-Arctic biodiversity of marine pelagic and sea-ice unicellular eukaryotes: A first attempt assessment. *Marine Biodiversity, 41*, 13-28.

Poulin, M., Underwood, G. J. C., and Michel, C. (2014). Sub-ice colonial Melosira arctica in Arctic first-year ice. *Diatom Research*, *29*(2), 213–221.

Protection of the Arctic Marine Environment (PAME). (2020). The increase in Arctic shipping: Arctic shipping status report (ASSR) #1. Retrieved from https://www.pame.is/projects/arctic-marine-shipping/arctic-shipping-status-reports/723-arctic-shipping-report-1-the-increase-in-arctic-shipping-2013-2019-pdf-version/file.

Rózańska, M., Gosselin, M., Poulin, M., Wiktor, J. M., and Michel, C. (2009). Influence of environmental factors on the development of bottom ice protist communities during the winter-spring transition. *Marine Ecology Progress Series*, *386*, 43–59.

Schein, A., Scott, J. A., Mos, L., Hodson, P. V. (2009). Oil dispersion increases the apparent bioavailability and toxicity of diesel to rainbow trout (*Oncorhynchus mykiss*). *Environmental Toxicology and Chemistry, 28*(3), 595-602.

Scott, M. and Hansen, K. (2016, September 16). Sea Ice. *NASA Earth Observatory.* Retrieved from https://earthobservatory.nasa.gov/features/SeaIce.

Sonune, A., and Ghate, R. (2004). Developments in wastewater treatment methods. *Desalination, 167,* 55-63.

Søreide, J. E., Leu, E. V. A., Berge, J., Graeve, M., and Falk-Petersen, S. (2010). Timing of blooms, algal food quality and *Calanus glacialis* reproduction and growth in a changing Arctic. *Global Change Biology*, *16*(11), 3154–3163.

Statistics Canada. (2016). Population size and growth in Canada: Key results from the 2016 census. Retrieved from https://www150.statcan.gc.ca/n1/daily-quotidien/170208/dq170208a-eng.htm?HPA=1.

Stroski, K. M., Luong, K. H., Challis, J. K., Chaves-Barquero, L. G., Hanson, M. L., Wong, C. S. (2020). Wastewater sources of per- and polyfluorinated alkyl substances (PFAS) and pharmaceuticals in four Canadian Arctic communities. *Science of the Total Environment, 708,* 134494.

Thyrring, J., Tremblay, R., and Sejr, M. K. (2017). Importance of ice algae and pelagic phytoplankton as food sources revealed by fatty acid trophic markers in a keystone species (Mytilus trossulus) from the High Arctic, *Marine Ecology Progress Series, 572*, 155–164.

United States Department of the Interior. (2008). Circum-Arctic Resource Appraisal: Estimates of Undiscovered Oil and Gas North of the Arctic Circle.

United States Environmental Protection Agency. (2012). Ecological Effects Test Guidelines OCSPP 850.4500: Algal Toxicity.

Von Quillfeldt, C. H., Ambrose, W. G., and Clough, L. M. (2003). High number of diatom species in first-year ice from the Chukchi Sea. *Polar Biology*, *26*(12), 806–818.

Wassman, P., Duarte, C. M., Agustí, S., Sejr, M., K. (2010). Footprints of climate change in the Arctic marine ecosystem. *Global Change Biology, 17*(2), 1235-1249.

Watanabe, H., Tamura, I., Abe, R., Takanobu, H., Nakamura, A., Suzuki, T, Hirose, A., Nishimura, T. and Tatarazako, N. (2015). Chronic toxicity of an environmentally-relevant mixture of pharmaceuticals to three aquatic organisms (alga, daphnid, and fish). *Environmental Toxicology and Chemistry, 35*(4),: 996-1006.

Weigel, S., Berger, U., Jensen, E., Kallenborn, R., Thoresen, H., and Hühnerfuss, H. (2004). Determination of selected pharmaceuticals and caffeine in sewage and seawater from Tromsø/Norway with emphasis on ibuprofen and its metabolites. *Chemosphere*, *56*(6), 583–592.

Weiss, C. M. (1976). Evaluation of the algal assay procedure. US Environmental Protection Agency. EPA-600/3-76-064.

2 A Critical Review of the Availability, Reliability, and Ecological Relevance of Arctic Species Toxicity Tests for Use in Environmental Risk Assessment

2.1 Introduction

Environmental risk assessors and others who rely on the ecotoxicological peer reviewed literature to aid in regulatory decision-making have identified data reliability and reporting as a critical issue preventing data uptake (see Borgert et al., 2016; Hanson et al., 2017; Harris and Sumpter, 2015; Thoré et al., 2019). In this study, reliability refers to the completeness of reported methodologies; specifically, whether the experimental design is sufficiently described to ensure reproducibility of the method and accuracy of the results (Ågerstrand et al., 2011a; Moermond et al., 2016) with the specific purpose of informing ERA. Well-designed studies with relevant species and endpoints contribute to the effectiveness of the ERA process (Ågerstrand et al., 2011a; Ågerstrand et al., 2011b; Küster et al., 2009). High quality data and effective reporting reduce the uncertainty around the potential for effects and gives risk assessors greater confidence in their decisions and recommendations. Furthermore, methodologically robust toxicity tests are vital in helping to reduce unnecessary animal testing within ecotoxicology (Burden et al., 2020; Vergauwen, 2018). Ultimately, well-designed and performed studies free-up resources that could be used to perform additional studies that will further ERAs, instead of repeating studies to confirm or refute inconsistent observations (Hanson and Brain, 2021).

While many regulatory agencies have internal screening criteria for data inclusion in ERAs (e.g., Henry and Pease, 2016; US EPA, 2003) there is currently no single, standard approach used to critically evaluate the reliability of data being used in ERA. Several approaches have been developed and proposed for ad hoc screening of the peer-reviewed literature (Ågerstrand et al., 2013; Durda and Preziosi, 2000; Hanson et al., 2019; Hobbs et al., 2005; Klimisch et al., 1997; Moermond et al., 2016; Schneider et al., 2009; Van Der Kraak et al., 2014). For example, Klimisch et al. (1997) proposed an evaluation system whereby data are sorted into one of four reliability categories for use in risk assessment ranging from reliable without restriction to not reliable or assignable. Criteria regarding thorough reporting of methods, results, and inclusion of appropriate

controls are outlined to aid in determining which category is appropriate for a particular study. This approach relies heavily on expert judgement, and thus, could be seen as relatively subjective in its application. Since then, several papers (e.g., Ågerstrand et al., 2013; Durda and Preziosi, 2000; Schneider et al., 2009) have expanded upon the relatively brief descriptors outlined by Klimisch et al. (1997) and broadened the scope beyond acute studies. Schneider et al. (2009) offered an instrument called ToxRTool whereby scores are given to different groups of criteria resulting in the assignment of the study to one of the Klimisch categories described above. This technique helps address the criticisms of subjectivity and lack of transparency that have been identified when using Klimisch categories as a stand-alone approach in determining data reliability (e.g., Kase et al., 2016). Ultimately, attempts to objectively assess data quality rely on a combination of using the above tools and expert judgement. Beyond assessing the quality of the data, a reasonable and neutral interpretation of results is crucial to ensure that ecotoxicology as a discipline maintains credibility in both the scientific and the public spheres (Hanson and Brain, 2021). Overemphasis of effects can not only have damaging implications for the reputation of the discipline, but can also tie up funding and resources in addressing problems that potentially do not require resolution (Hanson and Brain, 2020).

It is important to note that there are numerous reasons why a study may not be recommended for risk assessment purposes and these are not a judgment that the research is wholly uninformative or does not contribute to our larger understanding of risk. Experimental design and execution could be excellent, but if the question being asked is not relevant to a risk scenario, the utility of the data for ERA is likely limited. Relevance in this context refers to the applicability of the observations to ERAs and the utility of the data in informing decision-making (Ågerstrand et al., 2011a; Moermond et al., 2016). Whether specific data are relevant to a risk scenario depends on the risk assessment in question, but the ecological relevance of various endpoints (i.e., mortality being more definitively tied to population-level effects than biomarker responses) can be assessed independently of a specific risk scenario and was a focus of this review.

Thorough documentation and transparency are needed to accurately assess the reliability of data, but these elements can be lacking in published studies, leading to

difficulty in verifying the methods used and the results. Regulators often rely on studies conducted under Good Laboratory Practices (GLP) or according to standard methods (e.g., International Organization for Standardization [ISO], Organisation for Economic Co-operation and Development [OECD], United States Environmental Protection Agency [USEPA]), especially for plant protection products, as these guidelines have been developed to reduce uncertainty in toxicological testing (Borgert et al., 2016). However, it is not always possible to obtain these types of data for all compounds and other published literature is often the source for deriving acceptable environmental values (Ågerstrand et al., 2013; Henry and Pease, 2016). This is especially true for tests and assessments for ecological risk where there is a lack of standard test development, such as the Arctic.

It has been long noted in the literature that there is a lack of testing and understanding around Arctic species' sensitivity to contaminants and that this lack of information may result in undue harm to this ecosystem (Chapman and Riddle, 2003). Of the available data, only a fraction may be reliable enough to provide accurate insight into environmental risk. This gap is in part due to the lack of standardized laboratory testing of Arctic species as noted by others (Chapman and Riddle, 2005a). The response of relatively numerous temperate aquatic species to a vast array of contaminants has been characterized (Chapman and Riddle, 2003), but critical differences in the structure and function of Arctic and temperate ecosystems mean extrapolating these results could be a source of significant uncertainty. For example, Arctic ecosystems lack the same functional redundancy that is present elsewhere as there may only be one or two different species that play a similar functional role; while in temperate, tropical, or boreal ecosystems, there may be tens or hundreds of species with similar functional niches (Aune et al., 2018). Therefore, assessing and supporting the conduct of high-quality Arctic ecotoxicology data are vital to Arctic ecosystem protection.

The presence of contaminants in Arctic marine ecosystems (Sonne et al., 2021) underlines the importance of generating high quality, Arctic-specific toxicity data in order to provide accurate and robust insight for ERA in the region. Human activities in this region such as mineral extraction, hydrocarbon exploration, tourism, and shipping all

pose potential threats to flora and fauna, e.g., through the release of contaminants into aquatic ecosystems (AMAP, 2017; Aulanier et al., 2017; Hsiao, 1978). For example, in 2008, the United States Geological Survey estimated that up to 90 billion barrels of oil (~14 trillion litres), 1,669 trillion cubic feet (~47 trillion m^3) of natural gas, and 44 billion barrels (~7 trillion litres) of natural gas liquids remain undiscovered in the Arctic (United States Department of the Interior, 2008). If and when these resources are extracted, these processes bring to light several contaminants of concern; namely PAHs that are present in crude oil. As risk assessments move forward in the Arctic, a baseline of data gaps is needed to ensure proper test prioritization.

The objective of this critical review was to quantitatively characterize the current state of knowledge in Arctic ecotoxicological research regarding effects of contaminants on Arctic organisms. Specifically, we sought to objectively and transparently assess the methodological strengths and weaknesses (i.e., reliability) of available toxicity data found in the peer-reviewed and accessible grey literature through the creation and application of a scoring rubric. The results from the scoring rubric were used to analyze trends in current research practices, identify gaps in study design, conduct, and reporting, and to highlight the future work needed to address these deficiencies and contribute to the broader discussion of data reliability, both in Arctic ecotoxicology and beyond. Finally, this approach allowed for the highlighting of knowledge gaps in available Arctic species, endpoints, and contaminants with recommendations to address in future toxicity testing.

2.2 Methods

2.2.1 Literature Search and Initial Screening

An overview of the methodological process for this exercise is presented in Figure 2.1. Literature searches were conducted from January 2018 to June 2021 to gather relevant studies for assessment in this review. The University of Manitoba Libraries' database search engines were used to conduct Boolean searches from the Web of Science, Wiley Online Library, TOXLINE, ProQuest, and ScienceDirect databases. A Boolean search was conducted using the search terms: *Arctic* OR *polar*

AND *ecotoxicology*, *toxicology*, *toxic**, *effects*, *contaminants*, *test*, *assay*, *exposure*, NOT *Antarctic*. These search terms generated relevant articles that were then subject to backward and forward searching, which entails reviewing relevant sources cited by the obtained article and sources that have cited the article itself (Brocke et al., 2009).

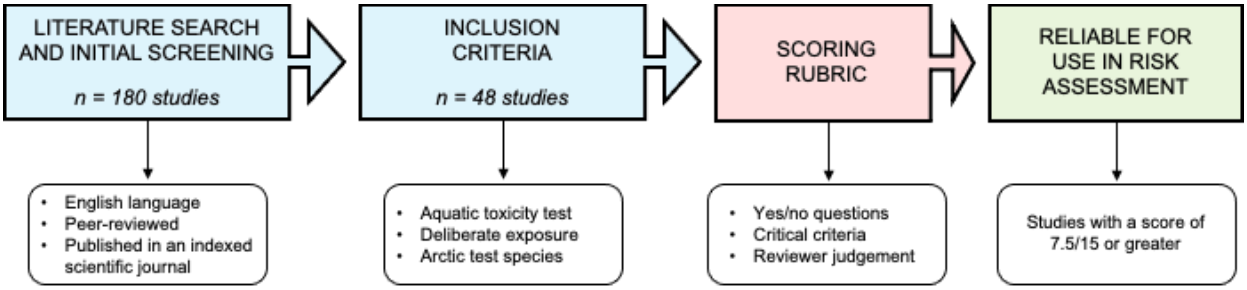

Figure 2.1 Overview of the process whereby Arctic laboratory toxicity tests were evaluated for the reliability and relevance of their data for use in environmental risk assessment.

Initial screening of retrieved articles was completed to filter out any non-English studies and those that were not peer-reviewed or published in an indexed scientific journal (i.e., one with an assigned impact factor). The process of backward and forward searching highlighted datasets available from conference proceedings and other grey literature (e.g., Arctic and Marine Oil Spill Program (AMOP), government technical reports) that were highly relevant to this subject matter. Despite targeting peer-reviewed publications for this exercise, the sporadic and sparse nature of Arctic species toxicity tests became apparent throughout the metadata collection phase. To address this, technical reports were included in the full evaluation (e.g., Dept. of the Environment Beaufort Sea Technical Report No. 11) where they were publicly accessible through database searches or easily available from a public forum. Where technical reports were not accessible (e.g., United States National Marine Fisheries Service technical reports), titles were gathered into a table as references (Table S1) but not scored. AMOP abstracts from 1978 through 2018 were accessed through personal copies available from the 2019 AMOP conference. At the time of this writing, AMOP abstracts were only available to the authors up until 2018. The inconsistency of including grey literature was outweighed by the importance of incorporating this widely accessible and

significant pool of Arctic-related data. It is also important to note that these publications represented a mechanism that was designed to complement standard peer-reviewed articles. A significant amount of Arctic ecotoxicology data has been gathered by Canadian government departments and programs (e.g., former Department of Indian Affairs and Northern Development, Newfoundland Oil Burn Experiment, Oil-in-Ice Joint Industry Program) where studies were not officially published or peer-reviewed, but excluding these data on this basis alone would have detracted from the overall purpose of this assessment and resulted in eight fewer studies from which data points could be gathered. For this reason, abstracts that reported results were subsequently included in the analysis if they also met the additional inclusion criteria outlined in the following section.

2.2.2 Inclusion and Exclusion Criteria

Reports and literature that passed initial screening were organized into a table where studies were further evaluated for their relevance to this review. The following criteria determined which studies proceeded to the scoring phase:

- Must be an aquatic toxicity test (freshwater, marine, or estuarine); terrestrial tests were excluded
- Must be laboratory toxicity data or deliberate field exposure (e.g., mesocosm studies; studies that measured effects in wild populations after accidental contaminant releases were excluded);
- Must use a test species endemic to an Arctic region, or be conducted in an Arctic context with a species that can be found in an Arctic region (see below for further details).

In summary, studies returned from the search were excluded if they were not a deliberate exposure to a test species found in the aquatic Arctic environment. Toxicity tests with seabirds were included as aquatic toxicity tests since seabirds spend more than 50% of their lives at sea (Brown, 1980). This behaviour means that they have the potential to be exposed to marine contaminants both directly (e.g., swimming through

contaminated areas) and indirectly (e.g., biomagnification via feeding on contaminated fish) (Braune et al., 2012; Campbell et al., 2005). Some retrieved studies included test species with Arctic distribution but that were not endemic to this region. For example, the blue mussel *Mytulis edulis* is commonly tested as a temperate species as it has geographical distribution throughout the North Atlantic ocean (e.g., Falfushynska et al., 2019); however, it has also been tested as an Arctic species due to its presence in the Arctic ocean (e.g., Thyrring et al., 2015). Similarly, the copepod *Calanus finmarchicus* exists geographically in both oceans and has been tested as both an Arctic (e.g., Faksness et al., 2011; Grenvald et al., 2013) and non-Arctic (e.g., Hansen et al., 2013; Øverjordet et al., 2014) species for comparison. In the case of the blue mussel, the study was included if conducted in an Arctic context (e.g., based on study objective, environmental parameters, or exposure scenario), but was excluded if conducted in a non-Arctic context. In the case of the copepod, *C. finmarchicus* was included from all papers as its prevalence and importance in Arctic regions are key to ecosystem processes (Hansen et al., 2009). Including all aquatic toxicity tests using a specific species, solely because they are found in Arctic regions although found elsewhere, would broaden the scope of this review beyond its original purpose of describing the current state of knowledge of Arctic-specific species within ecotoxicology.

2.2.3 Scoring

The objective of the scoring exercise was to assess the methodological strengths and weaknesses of published studies to determine the reliability of the data for use in ERA. To accomplish this, a scoring system was modified from Hanson et al. (2019), which was developed for primary producer toxicity tests and was based on previously published data reliability guidelines and criteria from standard methods (Klimisch et al., 1997; Schneider et al., 2009; US EPA, 2012). The aim of identifying scoring criteria *a priori* was to increase the transparency and objectivity of the process.
Information regarding the exposure scenario, test conditions, and results for each study (i.e., exposure durations, test vessels, environmental parameters, ECx values, statistical analyses performed) were incorporated into a spreadsheet to facilitate scoring and comparison between studies. The spreadsheet required modifications over several

iterations to ensure that pertinent information from all assessed studies could be captured (e.g., the inclusion of additional columns for media properties), but the basic scoring approach and weightings did not change. The system was comprised of three groups of criteria based on: 1) test substance; 2) test organism and experimental system; and 3) test design, statistics, and results. Within these groups, several key criteria were identified as critical study components (described in more detail in Section 2.4 below), and overall study scores were multiplied by a factor of 0.5 for each critical criterion that was not met (i.e., where the uncertainty introduced by its absence is considered exceptional). Experimental procedures were numerically scored with binary responses according to 15 criteria (i.e., 1 if the criterion was met and 0 if it was not). An additional category called “reviewer judgement” was added as part of the criteria when the method was outlined *a priori*, in the case that an unforeseen issue not captured in the rubric rendered a study unreliable (e.g., control performance was reported, but would not be considered adequate; or significant contamination of test compound in controls). This criterion ultimately did not have to be used for any of the studies assessed in this review. The scoring rubric is presented in Table 2.1. To be considered reliable for use in risk assessment, we set a minimum overall reliability score threshold of 7.5 out of 15, with all critical study criteria (those in bold, red typeface in Table 2.1) being met. Tests that achieved exactly 50% were considered to be reliable. Although the numerical value of 7.5 is arbitrary, an overall reliability score above this value means that the study did not achieve a score of 0 in any of the critical criteria which was our minimum threshold for being considered reliable.

Each endpoint that was assessed in each study was additionally given a score based on relevance of the response, which was not a factor in the overall reliability score, but was assessed separately to gauge utility for risk assessment. Endpoints were classified into one of eight categories based on the general descriptor of the response (i.e., biochemical, cellular, physiological, behavioural, growth, mortality, and reproduction, ecosystem-level) and scored based on potential impacts with increasing levels of biological organization. For example, endpoints that have more relevance to sustaining populations or communities are scored higher, while those with little or no linkage to higher-level effects were scored lower (US EPA, 2016). Though sublethal

endpoints such as behaviour and biomarker response are important to characterize and add value to our understanding of toxicity, they do not supersede endpoints that have more direct and definitive population-level effects such as mortality and reproduction. From a different perspective, data collected from tests that use lethality-based endpoints could be viewed as less useful due to the irreversible nature of death; for example, enforcing a regulatory threshold concentration of a contaminant that causes death in 50% of a population may result in too much harm already being done to the ecosystem in question as half of the organisms will have died. However, these EC/LC50 data are used to create species sensitivity distributions (SSDs) from which predicted no effect concentrations (PNECs) and concentrations considered hazardous for 5% of species (HC5s) are calculated (Wang et al., 2008). This process allows risk assessors to determine whether the level of risk to the ecosystem as a whole is acceptable, as some amount of risk is inevitable to allow human activities to take place. This endpoint weighting approach differs from current Canadian regulatory pass or fail practices but has been integrated into assessments such as the United States National Resource Damage Assessment (Office of the Federal Register, National Archives and Records Administration, 2021). A complete list of endpoints assessed in this study and justification for their inclusion in each category are presented in Table 2.2.

The rubric for the scored studies can be found in the Supplemental Information appended to this thesis. Other researchers are encouraged to modify criteria depending on their specific use and recalculate scores as needed if certain aspects of the sheet are more or less relevant to their particular scenario. A blank template is also included where criteria can be removed, added, or modified and used for additional studies.

Table 2.1 Scoring criteria, the rationale for its use, and elaboration on scoring used to assess the methodological quality of available Arctic species ecotoxicological data. Criteria presented in bold, red typeface are considered critical components of a study for reliable use in risk assessment. Each red criterion scored with a zero results in a multiplication factor of 0.5 on the overall reliability score of the study.

Group A – Test Substance				
	Criterion	***Rationale***	***Score of 1***	***Score of 0***
1	> 95% pure (or equivalent)	Using a test substance of high purity ensures that any toxic effects observed in the test are attributable to the active ingredient instead of other compounds in a formulated product.	If the source and percentage of the test substance were reported and the percentage was greater than 95% OR *the dispersant/oil type was identified*	If the source and/or percentage of the test substance were not reported or the percentage was less than 95% OR *the dispersant/oil type was not identified*
2	**Measured concentrations reported (any)**	**Analytically confirming tested concentrations rules out the possibility of test solution preparation errors or inconsistencies, giving greater confidence in the results.**	**If concentrations of any stock and/or test solutions were analytically confirmed and concentrations reported**	**If concentrations of any stock and/or test solutions were not analytically confirmed and no measured concentrations are reported**
3	Measured concentrations reported – individual initial	Analytically confirming tested concentrations rules out the possibility of test solution preparation errors or inconsistencies, giving greater confidence in the results; additionally, measuring concentrations at the start of the exposure provides an initial baseline against which later measurements can be compared.	If concentrations of test solutions in individual test units at test initiation (pooled or separate) were analytically confirmed and concentrations reported	If concentrations of test solutions in individual test units at test initiation (pooled or separate) were not analytically confirmed and no measured concentrations reported

Group A – Test Substance				
	Criterion	***Rationale***	***Score of 1***	***Score of 0***
4	Measured concentrations reported – individual final	Analytically confirming tested concentrations rules out the possibility of test solution preparation errors or inconsistencies, giving greater confidence in the results; additionally, comparing concentrations at the end of the exposure to those at the start can demonstrate how the compound has changed over the duration of the exposure. This is particularly useful in oil spill research.	If concentrations of test solutions in individual test units at the end of the test (pooled or separate) were analytically confirmed and concentrations reported	If concentrations of test solutions in individual test units at the end of the test (pooled or separate) were not analytically confirmed and no measured concentrations reported
5	**Number of concentrations tested (excluding control)**	**A minimum of three concentrations are recommended to effectively plot a concentration-response curve with confidence.**	**If there were ≥ 3 test concentrations, excluding control**	**If there were < 3 test concentrations, excluding control**
6	Ecological relevance	Data collected in laboratory toxicity tests are more useful to risk assessors if concentrations that can be found in the environment are included.	If at least one of the tested concentrations was less than or equal to environmentally-relevant concentration for that substance (see Table S2)	If none of the tested concentrations were less than the environmentally-relevant concentration for that substance (see Table S2)

Group B – Test Organism and Test System				
	Criterion	***Rationale***	***Score of 1***	***Score of 0***
7	Strain or source identified	Specific locations of wild-collected species may provide insight regarding potential previous	If the provenance of the test organism was identified (wild-collected	If the provenance of the test organism was not identified (wild-collected

Group B – Test Organism and Test System				
Criterion		***Rationale***	***Score of 1***	***Score of 0***
		exposures. Laboratory strains, depending on the species, may have adapted over time to culture conditions. Describing this information can help guide the interpretation of results based on these occurrences.	location, laboratory strain ID)	location, laboratory strain ID)
8	Initial test organism characteristics described	Physical characteristics, especially size, can directly influence the toxicity of a given compound, and relative responses between tests.	If initial characteristics relevant to the species (e.g., size, density, mass, feeding protocols) were described	If no initial characteristics relevant to the species (e.g., size, density, mass, feeding protocols) were described
9	Standard protocol followed	Standard methods allow more transparency in the data collection process by thoroughly describing methods and how results should be interpreted. Following these protocols also more readily facilitates intra- and inter-laboratory comparisons of data.	If the procedures were based on standard protocols (e.g., EPA, OECD, ASTM, ISO) or modified from previous studies; deviations acceptable if described	If the procedures were not based on standard protocols (e.g., EPA, OECD, ASTM, ISO) or modified from previous studies; deviations acceptable if described
10	Test conditions	Environmental conditions appropriate for the test species ensures that any toxic effects observed are due to the presence of the test substance and not any other environmental factor.	If relevant environmental parameters specific to the test species and substance are identified (see Table S4)	If relevant environmental parameters specific to the test species and substance are not identified (see Table S4)

Group C – Test Design, Statistics, and Results				
	Criterion	***Rationale***	***Score of 1***	***Score of 0***
11	**Replication**	**To perform meaningful statistical analysis of results, a minimum of three replicates are recommended in order to assess variation and contrast with controls.**	**Number of replicates is ≥ 3**	**Number of replicates is < 3**
12	**Statistical methods described**	**Describing statistical methods allows other members of the scientific community to observe how results were derived and determine whether the statistical tests were appropriate for the data that were gathered.**	**If the statistical method to calculate results (i.e., NOEC/LOEC/ECx/LCx, statistical significance between treatments) were described and reported and appropriate controls were employed**	**If the statistical method to calculate results (i.e., NOEC/LOEC/ECx/LCx, statistical significance between treatments) were not described or reported and appropriate controls were not employed**
13	Concentration-response	The inclusion of a concentration-response curve allows users of the data to see how the full toxicity profile is characterized at the full range of concentrations used in the test.	If a concentration-response model and parameters were reported in graph or formulaic expression	If a concentration-response model and parameters were not reported
14	Raw values	Including the raw dataset increases the transparency of the study by demonstrating how well the test was performed and allowing other users of the data to perform different statistical approaches if required.	If raw values (average raw values acceptable, but not % of control) were reported in tables and/or figures (all or average with error)	If raw values (average raw values acceptable, but not % of control) were not reported in tables and/or figures
15	Control performance	All tests should have some form of control performance criteria to demonstrate that observed toxicity is a direct result of the compound	If control values were reported and the reported values meet the control performance requirements	If control values were not reported or the reported values did not meet the control

Group C – Test Design, Statistics, and Results			
Criterion	***Rationale***	***Score of 1***	***Score of 0***
	instead of being attributed to poor husbandry or other factors.		performance requirement

Table 2.2 Scores for relevance of laboratory responses of available Arctic ecotoxicological data. They are based on their linkage to effects on organisms at the population and community-level. Study number reported in parentheses corresponds to the alphabetically-ordered list of assessed studies in Table S1.

Score	General Descriptor	Rationale	Endpoints Observed in This Study				
			Mammal	Bird	Fish	Invertebrate	Primary Producer
0	No known linkage to survival, development, growth, and/or reproduction	If a response has no real or hypothetical linkage to higher-level effects, then it has little to no value in elucidating ecological risk.	• NA	• NA	• NA	• NA	• NA
1	Biomarker response such as genomic or metabolomic measures.	While informative from a mechanistic perspective, the relevance of these responses to population, community and ecosystem-level effects is considered very low. In many cases, the responses characterized are regular processes to detoxify or adapt to a transient	• NA	• NA	• Liver transcriptome changes (1) • Gene expression (CYP1A) (27)	• Gene expression (CYP330A1) (18) • Gene expression (GST mRNA) (18, 19, 21, 39)	• Phosphate flux (44) • Potential nitrification (44) • Potential oxygen consumption (44) • Silicate flux (44)

Score	General Descriptor	Rationale	Endpoints Observed in This Study				
			Mammal	Bird	Fish	Invertebrate	Primary Producer
		stressor, which in and of themselves are not adverse.					
2	Biomarker responses such as enzymatic changes or general physiology.	While informative from a mechanistic perspective, the relevance of these responses to population, community and ecosystem-level effects is considered very low. In many cases, the responses characterized are regular processes to detoxify or adapt to a transient stressor, which in and of themselves are not adverse.	• Cell viability (6a, 6b, 14a, 14b) • Electrolyte changes (8) • Enzyme and metabolite changes (8) • Erythrocytic changes (8) • Hormone changes (8) • Intracellular thiol levels (14a, 14b) • Leukocytic changes (8) • Lymphocyte proliferation (6a, 6b, 14a, 14b, 29a, 29b, 29c, 29d) • Natural killer cell activity (6a, 6b) • Nitrogenous compound changes (8) • Phagocytosis (6a)	• NA	• Bile content/ analysis (7a, 7b) • EROD activity (7a, 7b, 16, 40) • Osmolality (40) • Plasmatic sex steroid measurements (16) • Plasmatic vitellogenin (16) • PAH metabolites (1, 16) • Thyroid status (40) • Vitamin status (40)	• Lysosomal membrane stability (10) • Acetylcholinesterase response (47) • Acyl-CoA oxidase response (47) • Glutathione S-transferase response (47)	• NA

Score	General Descriptor	Rationale	Endpoints Observed in This Study				
			Mammal	Bird	Fish	Invertebrate	Primary Producer
			• Metallothionein induction (14a, 14b)				
3	Changes in behaviour.	Behaviour, especially if related to reproduction, can have impacts on populations and communities of organisms. Other behavioural changes, such as predator avoidance could result in increased mortality.	• NA	• NA	• NA	• Attraction-repulsion response (42a) • Behavioural response (42a) • Curling of tube feet (9) • Emergence (9) • Feeding/feed response/grazing/ ingestion rate (42a, 46) • Foot protrusion (9) • Locomotory activity (43) • Loss of attachment ability (9) • Loss of cover (9) • Loss of equilibrium (45a, 45b) • Mantle gape (9) • Ostial closure (9) • Respiration (3b, 42a) • Retraction of tube feet (9) • Siphon retraction (9) • Spine droop/rigidity (9) • Stimulus response (9)	• NA
4	Changes in growth and development, such as mass,	These responses are typically highly relevant to the	• NA	• NA	• Cardiac activity and arrhythmia (34, 27)	• Fecal pellet production/volume (23, 25, 26, 28, 35, 46, 49)	• Chlorophyll-*a* content (28, 44)

Score	General Descriptor	Rationale	Endpoints Observed in This Study				
			Mammal	Bird	Fish	Invertebrate	Primary Producer
	other morphometrics, and phenology.	success or sustaining populations and communities in an ecological context.			• Condition index (7a, 7b) • Eye pigmentation (13a, 13b) • Time to hatch (13a, 13b) • Dry mass (27) • Standard length (27) • Myotome height (27) • Eye diameter (27) • Pericardial edema (27) • Lipid content (27)	• Lipid content (49, 50, 51) • Carbon content (50) • Nitrogen content (50) • Prosome length (50) • Fatty acid composition (48)	• Growth inhibition (4, 24) • ^{14}C incorporation (44)
5	Changes in reproduction, including malformation of young.	These responses are typically highly relevant to the success or sustaining populations and communities in an ecological context.	• NA	• NA	• Malformations (34, 27) • Underdeveloped embryos (13a, 13b)	• Clutch number (51) • Clutch size (51) • Development (4) • Egg production (23, 25, 26, 35, 38, 46, 49, 51) • Fraction reproducing (38) • Naupliar development (17) • Spawning interval (51)	• NA
6	Mortality and other surrogates.	Loss of individuals is highly relevant to the success or sustaining populations and communities in	• NA	• Survival to 90% of development (2)	• Mortality (1, 4, 5a, 5b, 11a, 11b, 13a, 13b, 15a, 27, 34, 37, 42a) • Hatching (13a, 13b, 27, 33)	• Mortality (3a, 3b, 3c, 3d, 3e, 4, 5a, 5b, 9, 10, 11a, 11b, 15a, 15b, 17, 18, 19, 20a, 20b, 20c, 21, 22a, 22b, 30, 31, 32, 33, 36, 37, 38, 39, 48,	• Biomass-specific primary production (BPP) (28) • Primary producer biomass

Score	General Descriptor	Rationale	Endpoints Observed in This Study				
			Mammal	Bird	Fish	Invertebrate	Primary Producer
		an ecological context.				49, 51, 42a, 45a, 45b, 46) • Hatching (17, 23, 25, 26, 38, 51)	(41a, 41b, 41c)

2.2.4 Data Handling

Metadata from this exercise were compiled into a spreadsheet for further consideration. The *in vitro* and *in vivo* studies followed the same scoring process, and in cases where a cell line was used, the test species and aquatic habitat (freshwater or marine) were categorized based on the source animal (e.g., ringed seal [*Pusa hispida*] cell tests filed under marine mammal tests). Unidentified test species (e.g., a sediment sample with unidentified bacterial and algal communities in Petersen et al. 2008) were counted as a discrete species category, as well as tests that identified the species only by genus (e.g., *Gammarus sp.* was differentiated from *Gammarus oceanicus* and counted as a different species).

Data are presented for individual papers (peer-reviewed publications (e.g., Olsen et al., 2013a), abstracts (e.g., Faksness et al., 2011), or theses [e.g., Moore, 2016]) collectively under the term 'studies', as well as by unique toxicity tests (test substance and species combinations, e.g., *Calanus glacialis* exposed to pyrene [Jensen et al., 2008]), or experimental combinations (test substance/species/endpoint combinations, e.g., ethoxyresorufin-O-deethylase [EROD] activity in Arctic cod exposed to produced water [Geraudie et al., 2014]). Test organisms were divided by organism group (i.e., invertebrate, vertebrate, or primary producer) and organism type as defined by the US EPA ECOTOXicology Knowledgebase (i.e., algae, bird, crustacean, fish, mammal, mollusc, or other invertebrate), with the additional category of microbial consortium included based on findings in the present study. Test substances were organized into three main groups based on their structure and use: oil-related contaminants, inorganic contaminants (e.g., metals), and organics (e.g., phenols and flame retardants). This was done in order to easily identify more generally groups of compounds where testing was either strong or lacking. Lastly, endpoints were categorized into eight effect categories again based on the US EPA ECOTOX Knowledgebase; behavioural, biochemical, cellular, ecosystem, growth, mortality, physiological, or reproduction. This was done in order to assist in the assignment of relevance scores to individual endpoints.

Oil-related contaminants in this context are defined as any type or brand of fuel oil or gasoline, products designed to clean up or remediate sites contaminated with fuel

oil or gasoline, and single compounds that are found in fuel oil or gasoline when tested in the context of fuel oil or gasoline contamination. Specific test substances grouped within oil-related contaminants were difficult to categorize due to the extensive number of products (e.g., types of oil and oil-spill remediation technologies), preparation methods (e.g., oil slick, mechanically or chemically dispersed water accommodated fractions [WAFs], *in-situ* burn residues), and additives (e.g., chemical dispersants or shoreline washing agents) used in this testing. To facilitate data analysis and assist in the understanding of knowledge gaps, crude oil was counted as a single test substance regardless of oil type, preparation method, or additives. Oil-spill remediation technologies were counted as a single test substance based on operational usage (e.g., Absorrep K212 and Bioversal both counted as shoreline washing agents; Dasic NS and Corexit 9500 both counted as dispersants).

To investigate whether there was a correlation between the impact factor of the journal in which a paper was published and the reliability of the data presented, historical impact factors were retrieved from the Web of Science InCites Journal Citation Report (Clarivate Analytics, 2019) for the year in which the study was published. Where no impact factor was available for the time of publication (e.g., some articles published in the 1970s and 1980s), the average from the most recent five years was calculated and used in its place.

Although the scoring rubric was designed to be as objective as possible, there remained areas that require clarification for full transparency. For example, Criterion 11 requires a minimum of three replicates per exposure to achieve a score of 1; however, the line between true replication and pseudoreplication is sometimes difficult to discern (e.g., replicate dishes being held in the same bath in Frantzen et al., 2012 or the inclusion of twelve eggs per dose group pooled together for survival calculations in Braune et al., 2012). Arguments could be made that replicates using the same preparation of stock or test solutions are actually pseudoreplicates; or, similarly, EUs being incubated in the same growth chamber. There are different approaches to deal with pseudoreplication in study design such as nesting pseudoreplicates under genuine replicates (e.g., Lazic et al., 2020); however, to limit subjective differences in scoring

between studies for this criterion, replicates were simply defined as unique exposure vessels.

Environmental relevance of test substances (Criterion 6) is described in detail in Table S2. Note that environmentally-relevant concentration thresholds listed in this table are based on general literature values, but study-specific justification may change the environmentally-relevant value for a particular study. For example, a general environmentally-relevant concentration of PAHs may be 189 µg/L based on data accumulated from an oil spill; however, Petersen et al. (2008) used a measured value of sediment porewater concentrations to justify environmental realism as their study mirrored this type of exposure, and as such were scored a 1 in this category.

The purity and/or grade of some oil-related contaminants are difficult to characterize and compare across studies (Criterion 1). Justification for purity or grade of these test substances can be found in Table S3. Similarly, relevant environmental parameters to include when describing methodologies for toxicity tests can vary depending on test species and whether the test was conducted *in vitro* or *in vivo*. Justification for scoring in this criterion can be found in Table S4.

2.2.5 Statistical Analyses

For comparisons, scores from each group of criteria (test substance; test organism and system; test design, statistics, and results) were converted to percentages of the total achievable score in each group. These percentages were used for statistical analyses. Before assessing the data for statistically significant differences between organism groups, assumptions of normality were assessed using the Shapiro-Wilk test and homogeneity of variances using Bartlett's test (base package in R; R Core Team, 2012). Then, statistical significance of differences between organism groups were assessed through Kruskal-Wallis rank tests. Post-hoc Dunn tests were performed with a Bonferroni *p* adjustment to reveal where significant differences occurred. Spearman's correlation coefficients (rho) were calculated for non-normally distributed data with unequal variance using linear regression models. Scores are presented in the results section of this paper as mean ± standard deviation.

2.3 Results

2.3.1 Summary of Reviewed Studies

The search process returned 180 studies, 47 of which passed screening criteria for scoring. One additional PhD thesis (Moore, 2016) was included as it was publicly available and included a large volume of test results from freshwater Arctic fish. This addition resulted in a total of 48 individual studies (40 peer-review and 8 grey literature). The four data chapters of the thesis were assessed independently, resulting in a total of 51 separate scoring sheets in the Excel file (See Supplemental Information – Study Scores). There were 253 unique toxicity tests (i.e., test substance and test species combinations), 73 unique Arctic test species (two unidentified groups of microbes not included in this total; Pančić et al., 2019 and Petersen et al., 2008), 39 tested substances, and 95 endpoints, for a total of 596 experimental combinations (i.e., test substance/test species/endpoint combinations).

The volume of research produced in this field steadily declined from the 1970s (*n* = 45 unique toxicity tests) through the 1980s (*n* = 18) and 1990s (*n* = 0), and began to increase again in the 2000s (*n* = 5), finally experiencing a period of rapid growth in the 2010s (*n* = 185); however, 104 tests of the 185 reported in the 2010s were from a single study (Moore, 2016). Excluding this study, 81 tests were performed in this decade. The linear relationship between number of unique toxicity tests conducted and study year was not statistically significant (Figure S1). Furthermore, the thesis of Moore (2016) was the driver of a notable increase in testing of inorganic contaminants in the 2010s (Figure S2). When excluded, oil-related contaminants comprised the majority of tests throughout this decade.

2.3.2 Test Species

Vertebrates were tested in 141, invertebrates in 104, and primary producers in 8 out of 253 unique toxicity tests (Figure S3). These organism groups were further broken down by organism type within each category (Table 2.3). Of these, 107 of 253 discrete experiments used freshwater organisms (42% of all unique tested species) and 146 of 253 used marine organisms (58%; Figure S4). Crustaceans were the most studied

organism type in number of individual studies (*n* = 29, 60% of all studies), number of species tested (*n* = 31, 42% of all species), and number of endpoints studied (*n* = 26, 27% of total endpoints). They were also one of only two organism groups that encompassed both marine (*n* = 29 species) and freshwater (*n* = 2 species) organisms. Within this group, the class Malacostraca were the most well-studied (*n* = 41, 47% of crustacean tests) and the most diverse class in terms of number of species tested (*n* = 20, 23% of species within crustaceans). Hexanauplia were the second-most well studied class of crustaceans (*n* = 34, 39% of crustacean tests), and within Hexanauplia, *Calanus glacialis* (*n* = 15, 44% of Hexanauplia tests) and *Calanus finmarchicus* (*n* = 14, 41% of Hexanauplia tests) were the primarily studied species. These two *Calanus* copepods were also the most frequently tested organisms of all crustaceans, and represented the only subclass of Hexanauplia that was tested (Copepoda; no organisms from the subclass Tantulocarida were tested).

Table 2.3 Summary of publicly available Arctic ecotoxicological laboratory tests that passed inclusion criteria for the present study as of June 2021.

Organism Type	**# Individual Studies *(n = 48)***	**# Unique Toxicity Tests *(n = 253)***	**# Species Tested *(n = 73)*[a]**	**# Compounds Tested *(n = 39)***	**# Endpoints Assessed *(n = 95)***
Crustacean	29	87	31	12	26
Fish	14	126	12	24	24
Mammal	4	13	8	8	12
Algae	3	6	6	3	3
Mollusc	5	12	11	5	13
Other Invertebrate	4	5	3	2	10
Microbial Consortium	2	2	NA	2	13
Bird	1	2	2	1	1

[a] Two unidentified groups of microbes not included.

Fish were the second most frequently tested organism type by number of individual studies (*n* = 14, 29% of studies) but were the most well-studied in terms of unique toxicity tests (*n* = 126, 50% of tests). This group was the second-most diverse in terms of number of species tested (*n* = 12, 16% of all tested species); however, fish

tests were entirely comprised of a single organism class (Actinopterygii, or ray-finned fishes). Both marine (*n* = 7 species) and freshwater species (*n* = 5 species, anadromous fish included) made up this group; however, the majority of data points collected using fish were driven by a single study that tested a variety of metals, metal salts, and nitrogenous contaminants on freshwater fish (Moore, 2016). With this study included, *Prosopium cylindraceum* (round whitefish; *n* = 25, 20% of fish tests) and *Salvelinus alpinus* (Arctic char; *n* = 25, 20%) were the most frequently tested species. With this study excluded, the keystone species, *Boreogadus saida* (known by the common names Arctic cod in Canada and polar cod in Europe) was the most frequently tested (*n* = 11, 50% of remaining fish tests). Fish were tested against the widest variety of test substances compared to other organism groups (*n* = 24 unique test substances, 62% of all tested compounds) representing all three contaminant groups (oil-related contaminants, inorganic contaminants, and organic contaminants).

Mammals were the third-most frequently studied organism types based on number of individual studies (*n* = 4, 8% of studies) and number of unique toxicity tests (*n* = 13, 5% of unique tests). Twelve of the thirteen mammalian tests were *in vitro* studies using orca (*Orcinus orca*), seal (*Pusa hispida, Pagophilus groenlandicus*), and polar bear (*Ursus maritimus*) cell lines to determine cellular, biochemical, and physiological effects. A single test (Engelhardt, 1981) was conducted using whole organisms to assess the clinical effects of exposure to oil pollution on polar bears. Microbial, algal, mollusc, other invertebrates, and bird tests were a combined total of only 11% of unique toxicity tests (*n* = 27 tests). Algal tests principally used Arctic diatoms as test species (i.e., Bacillariophyceae and Coscinodiscophyceae; *n* = 4, 67% of algal tests) but green algae were also tested (Chlorophyceae; *n* = 1, 17% of algal tests). A single test was conducted on an unidentified species of phytoplankton (Lemcke et al., 2018). The “other invertebrate” category is composed of invertebrates that did not fall under either of the other organism types (i.e, crustacean or mollusc). This category includes the green sea urchin (*Strongylocentrotus droebachiensis*), sea spider (*Nymphon gracile*), and balloon jellyfish (*Halitholus cirratus*). The microbial consortium category captured two studies (Pančić et al., 2019; Petersen et al., 2008) that used an unidentified mixture of benthic algae and bacteria obtained from grab samples as test

species. Seabirds were also tested in a single study (Braune et al., 2012). Scores from the studies testing seabirds and microbes were omitted in comparisons of scores between organism types to avoid skewing the data based on one and two studies, respectively, but were included in assessments of the dataset as a whole (e.g., included when presenting number of tested compounds and endpoints among all studies, but excluded in statistical analyses comparing overall reliability scores of different organism types).

2.3.3 Test Substances

Contaminants used as test substances in Arctic ecotoxicological research were categorized into three groups: oil-related contaminants, inorganic contaminants, and organic contaminants. The majority of experiments using inorganic contaminants were freshwater tests (*n* = 104 of 114 tests; 91%) while both other compound groups were mainly comprised of marine tests (98% of oil-related contaminant tests [*n* = 125 of 127 tests] and 92% of organic contaminant tests [*n* = 11 of 12 tests]); Figure S5). Oil-related contaminants were used in 50% of all toxicity tests (*n* = 127), with crude oil comprising the majority of substances in this group (*n* = 63, 50% of tests using oil-related contaminants). Crude oils from both Arctic (Alaska North Slope, Atkinson Point, Norman Wells, Cook Inlet, Kobbe, Troll, North Sea, and Prudhoe Bay) and non-Arctic regions (Venezuela, Arabian Light, Lago Medio, Midale, and Pembina) were included and were used as fresh and weathered, with and without suspended sediment, and in the presence and absence of ultraviolet light. Additionally, specific compound groups found in oil (e.g., aromatic, asphaltic or paraffinic compounds alone) were tested as individual test substances. Preparations of crude oil were either mechanically dispersed (i.e., physically agitated to create a WAF of oil) or chemically dispersed (i.e., crude oil used in conjunction with a chemical dispersant to create a chemically-enhanced water accommodated fraction [CEWAF] of crude oil), or used after a period of natural attenuation by microorganisms was allowed to occur, as *in-situ* burn residues, or as a slick in mesocosm studies. Single compounds found in oil (naphthalene, 2-methyl-naphthalene, and pyrene) were the second-most frequently tested group of substances in this category, used in 27% of oil-related contaminant tests (*n* = 34). Spill response

agents alone (i.e., no oil added) were used in 13% of oil-related contaminant testing using either chemical dispersants (Corexit [9500, 9500A, or 9527], Dasic NS, Finasol OSR52, or Gamlen OD4000) or shoreline washing agents (Absorrep K212, Bios, Bioversal, Corexit 9580A, or Hela saneringsvaeske). The remaining tests in this category were conducted using heavy fuel oil (n = 2) and produced water (n = 7), an oil extraction by-product. The vast majority of crustacean, algae, mollusc, other invertebrate, and microbial tests have been conducted with oil-related contaminants (Figure 2.2), but there are some data for these substances across nearly every organism category.

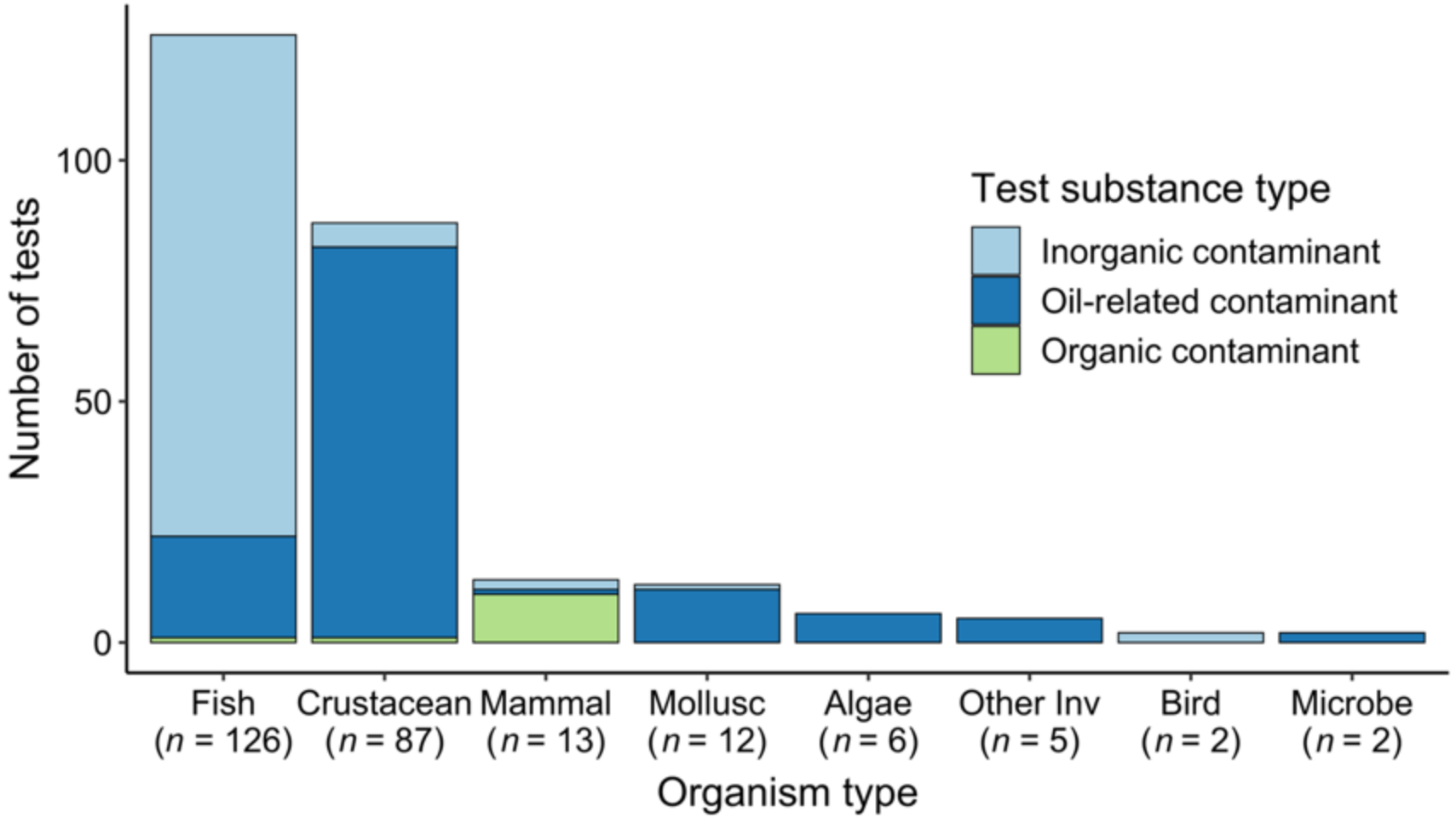

Figure 2.2 Number of unique Arctic species toxicity tests captured in this review (*n* = 253 as of June 2021) by organism group, with test substance type within bars (number below each bar is total number of unique tests).

Inorganic contaminants were made up of mostly metals and salts (e.g., copper, zinc, aluminum, sodium chloride, calcium chloride) and were the second-most tested class of substances, used in 45% of all tests (*n* = 114). More than three quarters (*n* = 104) of all fish tests in particular used inorganic contaminants, most of which were tested in the same study (Moore, 2016). They were also the only class of substances tested on seabirds. This compound group was the most diverse and included 23 unique test substances.

Organic contaminants were the least frequently tested compounds encompassing just under 5% of all toxicity tests (*n* = 12), all of which were tested on mammalian cell lines (*in vitro* studies) with the exception of a single test each with crustaceans and fish. Half of the tests that used this group of substances were from a single study (Desforges et al., 2017) where contaminant cocktails were derived from the blubber of wild marine mammals. Upon analysis, they were determined to be composed primarily of PCBs and organochlorine pesticides. The remaining compounds tested in this category were various PCB congeners, tetrachlorobiphenyl, and unidentified phenolic compounds from crude oil.

2.3.4 Endpoints

The 95 distinct endpoints were grouped into eight effect categories as shown in Figure 2.3. Mortality was the endpoint most frequently assessed with 54% of all data points being attributed to this effect category (*n* = 319 of 596 experimental combinations). Within this effect category, crustaceans comprised the majority with 42% being attributed to this group (*n* = 134, 42% of mortality endpoints), followed by fish (*n* = 128, 40% of mortality endpoints), molluscs (*n* = 9), other invertebrates (*n* = 4), and birds (*n* = 2). All other effect categories were sublethal, chronic, and/or ecosystem-based, and take up a combined total of less than half of all data points observed (*n* = 277, 46% of all data points). Behavioural endpoints were the second-most frequently assessed comprising 14% of all tested combinations (*n* = 85). Within this effect category, crustaceans were again the most popular test organism at 78% of all data points (*n* = 66), followed by other invertebrates (*n* = 11) and molluscs (*n* = 8). The third most frequently assessed endpoint was growth, which was the primary endpoint for algal

studies (*n* = 13 of 17 algal data points; 25% of all growth endpoints). Endpoints for algal growth inhibition assays were classified under “growth” rather than “reproduction”. This endpoint was also assessed in crustaceans (*n* = 19 of 51 growth endpoints; 37% of all growth endpoints), fish (*n* = 17; 33%), and molluscs (*n* = 2; 4%) as test species. The remaining five effect categories comprised a total of only 24% of all data points (*n* = 141). In order of most frequent appearance, they included biochemical endpoints (*n* = 50; 8% of all endpoints), cellular (*n* = 37; 6% of all endpoints), physiological (*n* = 28; 5% of all endpoints), reproduction (*n* = 20; 3% of all endpoints), and ecosystem-based responses (*n* = 6; 1% of all endpoints).

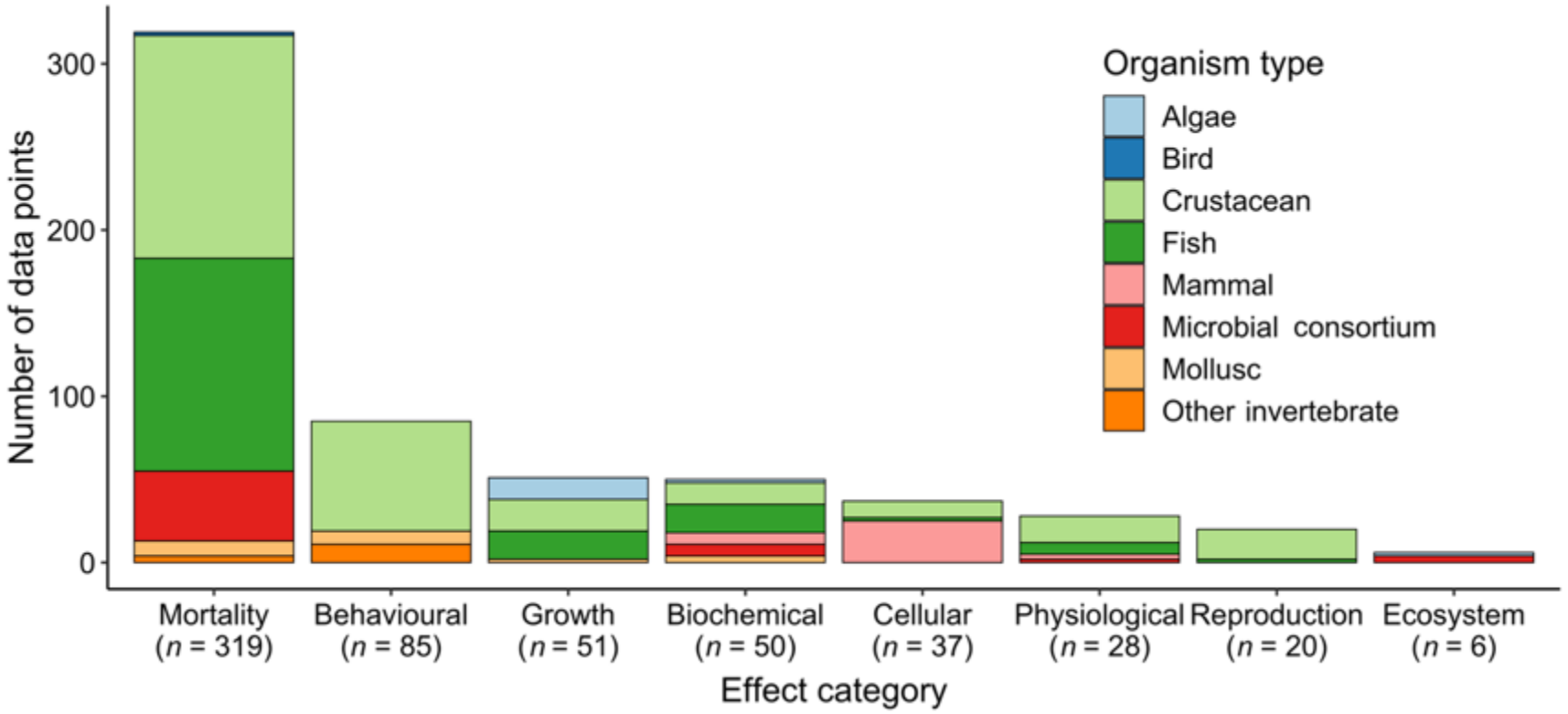

Figure 2.3 Number of data points for Arctic species captured in this review (*n* = 596 as of June 2021) by effect category, with organism type within bars (number below each bar is total number of reported responses for that category).

A period of monitoring post-exposure, framed as either a recovery period or opportunity to observe latent effects was included in 25% of individual studies (n = 12 of 48 studies), The terminology used in the study was mirrored in the scoring spreadsheet to present the results for these time-points in the same context as originally reported.

2.3.5 Overall Methodology Scores

Overall reliability scores calculated from the rubric are presented as vertebrates, invertebrates, or primary producers in Figure 2.4. Scores were variable within and between groups; however, the majority of means, medians, first, and third quartiles for every group were below 50% of the total achievable score. Mean scores expressed as a percentage of the total achievable score (i.e., divided by 15 and multiplied by 100) were greatest for vertebrates (36% ± 15) followed by invertebrates (30% ± 20) and primary producers (30% ± 23). The difference in means of overall reliability scores between vertebrates and invertebrates were found to be statistically significant, while primary producers were not significantly different than either of the other two groups ($p < 0.05$).

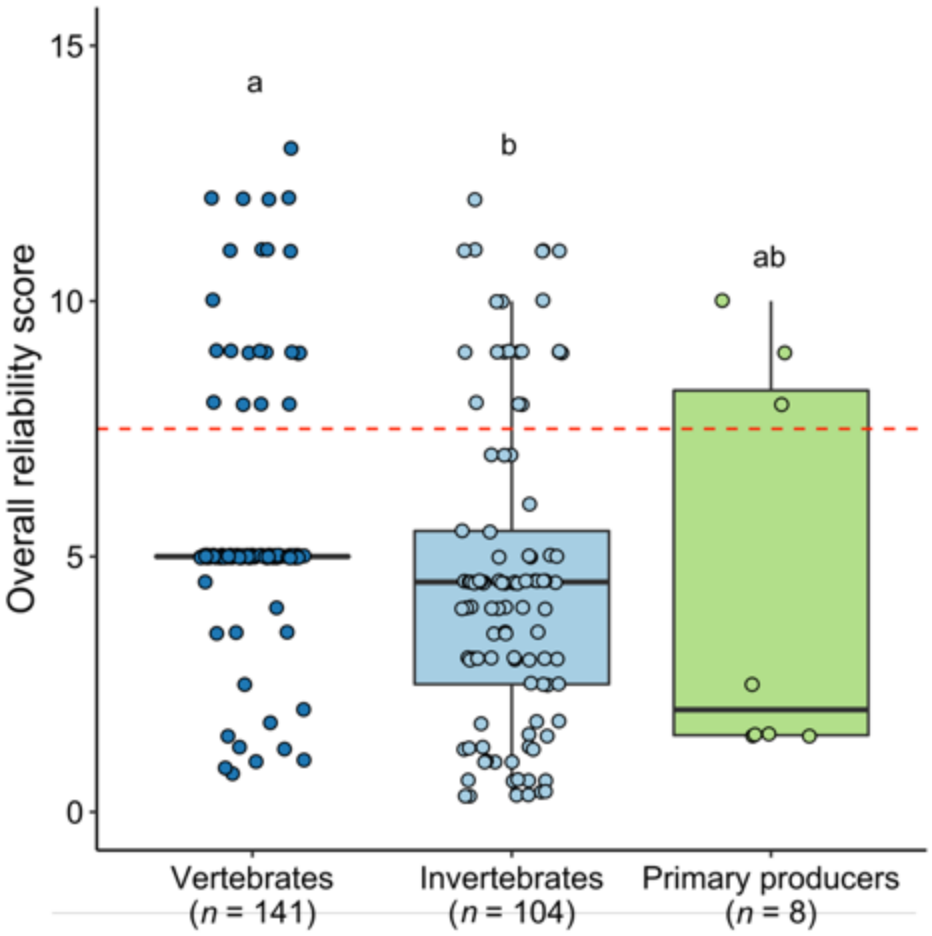

Figure 2.4 Overall reliability scores for all unique toxicity tests reviewed in this study (*n* = 253 as of June 2021) by their status as vertebrates, invertebrates, or primary producers (number below each bar is total number of unique toxicity tests for that group). Median values are represented by a black line within each box, lower and upper hinges correspond to the 25th and 75th percentiles, respectively, and whiskers extend to the largest or smallest value within 1.5x the interquartile range from the hinge. Letters indicate statistical significance between groups ($p < 0.05$) based on Kruskal-Wallis rank sum tests followed by post-hoc Dunn tests with Bonferroni *p* adjustments.

The individual toxicity test scores for groups A (test substance), B (test organism and test system), C (test design, statistics, and results), and overall are presented by organism type in Figure 2.5. Birds and microbes have been omitted from this figure and statistical analyses due to the small number of scores obtained from these groups of studies (n = 2 for each); however, mean scores for each criteria group for all organism types can be found in Table 2.4. Tests achieved the greatest score in Group B (74% ± 28), followed by Group A (55% ± 16), Group C (50% ± 15), and overall across all organism types (34% ± 18). The overall reliability scores were generally lower than the average scores for specific groups of criteria due to the high proportion of tests that did not meet one or more of the critical study criteria outlined in the rubric, resulting in a multiplication factor of 0.5 applied to the overall reliability score for each one not met.

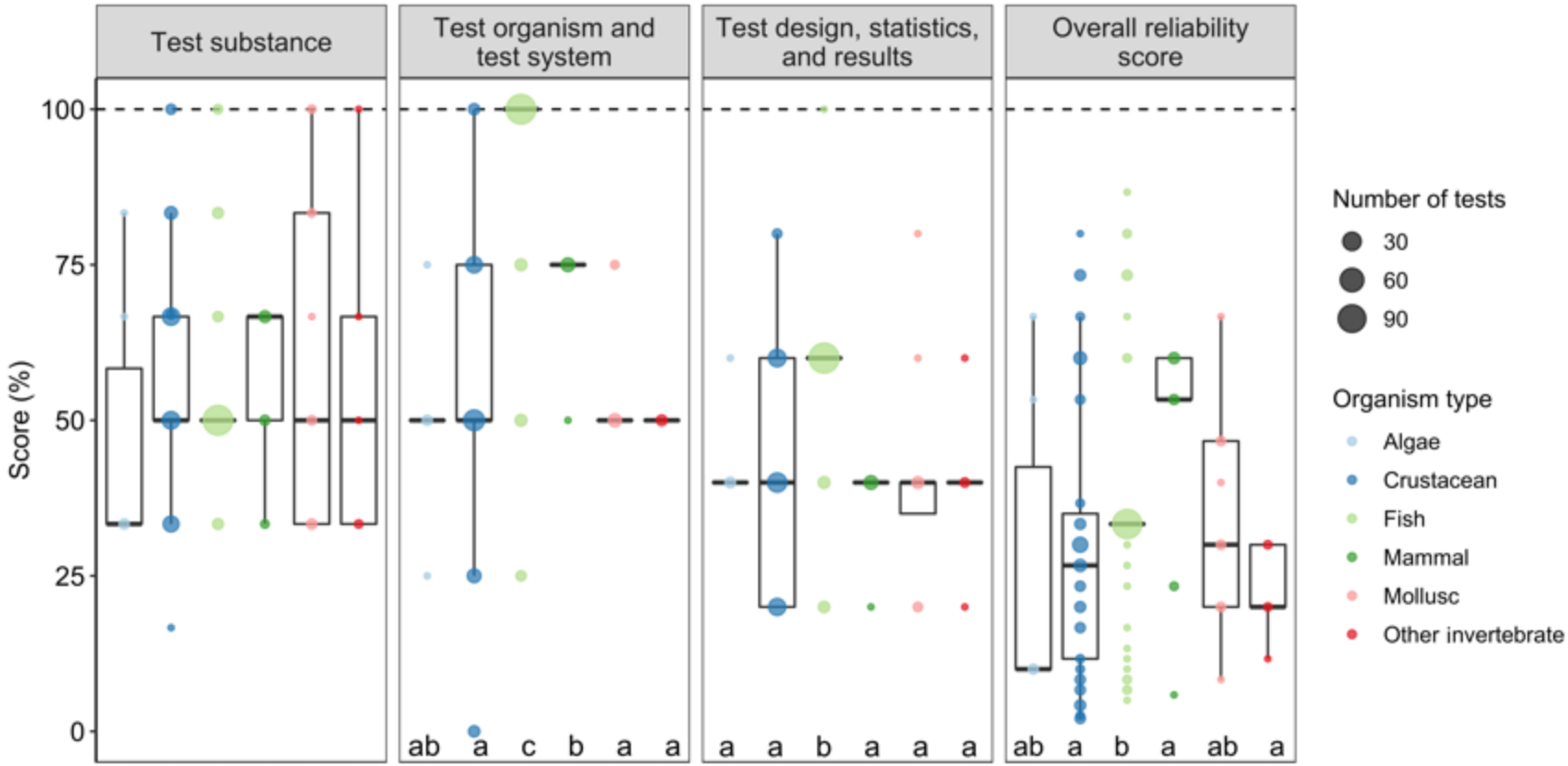

Figure 2.5 Scores (as a percentage out of 15) for all unique toxicity tests reviewed in this study (*n* = 253 as of June 2021) by organism type. Median values are represented by a black line within each box, lower and upper hinges correspond to the 25th and 75th percentiles, respectively, and whiskers extend to the largest or smallest value within 1.5x the interquartile range from the hinge. Letters indicate statistical significance between groups ($p < 0.05$) based on Kruskal-Wallis rank sum tests followed by post-hoc Dunn tests with Bonferroni p adjustments. No statistically significant differences were found for Group A (test substance)

Within Group A, toxicity tests performed on molluscs had the greatest mean score (60% ± 26), although other invertebrates, crustaceans, mammals, fish, and algae were closely behind with mean scores ranging between 47% and 57%. Tests performed with algae scored the lowest in this criteria group (47% ± 22). Variability in scores was the greatest in this category for other invertebrates (coefficient of variation [CV] = 49%) while fish scores were the most consistent (CV = 21%).

Tests performed with fish had the greatest mean score among organism types for Group B (93% ± 18), followed by mammals (73% ± 7), molluscs (54% ± 10), crustaceans (53% ± 24), other invertebrates (50% ± 0), and algae (50% ± 16). Variability in scores for criterion Group B was greatest for tests using crustaceans (CV = 46%) with scores ranging from zero to 100%. The least variability was found in the other invertebrate group (CV = 0%).

Within Group C, fish again had the greatest average score (57% ± 10). Algae, crustaceans, molluscs, and other invertebrates all scored in the 40% range (43% ± 8; 42% ± 17; 40% ± 17; 40% ± 14, respectively) and mammals achieved the lowest average score of 38% ± 6. Variation in scores was greatest for molluscs (CV = 43%) and least for mammals (CV = 14%).

With regard to overall reliability scores (i.e., the sum of groups A, B, and C multiplied by 0.5 for each critical criterion not met), mammalian tests scored the greatest with an average score of 48% ± 18) and also had the greatest percentage of unique toxicity tests that were deemed most reliable for use in risk assessment (e.g., score greater than 50%) according to the criteria outlined herein (*n* = 10 of 13 mammalian tests, 77%). Fish tests scored second-greatest overall at 35% ± 13) followed by molluscs (34% ± 16), crustaceans (30% ± 21), algae (27% ± 26), and other invertebrates (22% ± 8). After mammals, algal tests had the second-highest proportion of reliable studies with 33% of tests meeting the criteria, although only six unique toxicity tests were observed in total. Twenty of 87 crustacean tests were considered reliable for risk assessment (23%), 1 of 12 mollusc tests (8%), 10 of 126 fish tests (8%), and zero of five tests using other invertebrates. In total, only 46 of 253 (18%) unique toxicity tests passed screening for data reliability for use in ERA.

Finally, statistically significant co-relationships were found between score and study year for Groups B, C, and overall reliability scores, but not for Group A (Figures S6-9). The co-relationship between overall reliability score and the impact factor of the journal in which the study was published was analyzed and determined to be statistically significant (Figure S10).

Table 2.4 Percentage of unique toxicity tests with a score of 1 for each criterion by organism type. Bolded numbers indicate < 50% of tests with a score of 1 for that criterion. Mean reliability scores ± standard deviation for each organism type are calculated per criterion group and presented in italics below each section.

		Crustacean	Fish	Algae	Mammal	Mollusc	Other Inv.	Bird	Microbe	All Unique Tox Tests
Group A	Criterion 1	79	95	100	54	**33**	60	100	50	84
	Criterion 2	64	98	**33**	85	67	**40**	100	100	80
	Criterion 3	**15**	**7**	**17**	**0**	**33**	**20**	**0**	50	**11**
	Criterion 4	**14**	**5**	**0**	**8**	**33**	**20**	100	100	**11**
	Criterion 5	71	94	100	92	92	100	100	50	86
	Mean Score	*57 ± 19*	*52 ± 11*	*47 ± 22*	*56 ± 13*	*60 ± 26*	*57 ± 28*	*83 ± 0[a]*	*75 ± 12*	*55 ± 16*
Group B	Criterion 6	94	**16**	33	100	100	100	100	100	56
	Criterion 7	93	100	100	100	100	100	100	100	98
	Criterion 8	79	97	83	100	100	100	100	100	91
	Criterion 9	**15**	88	**17**	**0**	**8**	**0**	**0**	**0**	**49**
	Mean Score	*53 ± 24*	*93 ± 18*	*50 ± 16*	*73 ± 7*	*54 ± 10*	*50 ± 0[a]*	*75 ± 0[a]*	*63 ± 18*	*74 ± 28*
Group C	Criterion 10	**23**	89	**0**	92	**8**	**0**	100	50	57
	Criterion 11	55	**9**	100	92	67	**40**	100	50	**37**
	Criterion 12	84	100	**33**	92	100	100	100	100	93
	Criterion 13	**0**	**1**	**0**	**0**	**0**	**0**	**0**	50	**1**
	Criterion 14	**38**	90	83	**8**	**17**	**20**	100	50	61
	Criterion 15	**32**	87	**0**	**0**	**17**	**40**	100	**0**	55
	Mean Score	*42 ± 17*	*57 ± 10*	*43 ± 8*	*38 ± 6*	*40 ± 17*	*40 ± 14*	*80 ± 0[a]*	*50 ± 14*	*50 ± 15*
	Mean Reliability Score	**30 ± 21**	**35 ± 13**	**27 ± 26**	**48 ± 18**	**34 ± 16**	**22 ± 8**	**80 ± 0[a]**	**38 ± 31**	**34 ± 18**

[a] All scores identical

2.3.6 Individual Criteria Scores

Data pertaining to individual criteria organized by organism type are presented in Table 2.4 and are visually represented in Figures S11-13. Bolded, red numbers indicate that fewer than 50% of tests achieved a score of 1 in that criterion. Across all organism types, Criterion 7 (strain or source of test organism identified) was met the most frequently (*n* = 247 of 253 unique toxicity tests; 98%), followed by Criterion 12 (statistical methods described; *n* = 234, 92%), Criterion 8 (initial test organism characteristics described; *n* = 230, 91%), Criterion 5 (three or more test concentrations included; *n* = 217, 86%), Criterion 1 (source and purity of test substance reported; *n* = 212, 84%), Criterion 2 (any analytical confirmation of tested concentrations; *n* = 206, 81%), Criterion 14 (key measured raw values included; *n* = 159, 63%), Criterion 10 (relevant environmental parameters identified and reported; *n* = 148, 58%), Criterion 15 (control values reported and performance criteria met; *n* = 143, 57%), Criterion 6 (at least one exposure concentration < environmental concentrations; *n* = 138, 55%), Criterion 9 (standard protocol followed; *n* = 126, 50%), Criterion 11 (minimum of three replicates per exposure; *n* = 90, 36%), Criterion 3 (individual initial analytical confirmation of tested concentrations; *n* = 29, 11%), Criterion 4 (individual final analytical confirmation of tested concentrations *n* = 28, 11%), and Criterion 13 (concentration-response model and parameters provided; *n* = 2, 1%).

Several organism types achieved a score of 1 for specific criteria in 100% of unique toxicity tests. These were algal tests in Criteria 1, 5, 7, and 11; fish tests in Criteria 7 and 12; mammalian tests in criteria 6, 7, and 8; mollusc tests in Criteria 6, 7, 8, and 12; and other invertebrate tests in Criteria 5, 6, 7, 8, and 12. Additionally, almost all fish tests achieved a score of 1 in Criteria 2 and 8 (98% and 97%, respectively). In contrast, several organism types did not achieve a score of 1 in any test for certain criteria. No algal tests achieved a score of 1 in Criteria 4, 10, 13, or 15; likewise for crustacean tests in Criterion 13; mammal tests in Criteria 3, 9, 13, or 15; mollusc tests in Criterion 13; and other invertebrate tests in Criteria 9, 10, or 13.

Figure S14 shows overall, among all organism types, the percentage of unique toxicity tests that achieved a score of 1 in the four critical criteria outlined in the rubric (criteria in

bolded, red text in Table 2.1). Of these four criteria considered critical components of a laboratory toxicity test for reliable use in risk assessment, the inclusion of at least three replicates (Criterion 11) was met in the fewest number of unique toxicity tests (*n* = 90, 36%). Eighty percent of tests (*n* = 206) included some analytical confirmation of test concentrations (i.e., measured concentrations in stock solution prior to diluting to create test solutions, or in test solution stocks prior to allocating into individual EUs; criterion 2). The inclusion of a minimum of three test concentrations (Criterion 5) was met in 86% of tests (*n* = 217), and an appropriate description of statistical methods (Criterion 12) was met in 92% of tests (*n* = 234). These data are further broken down by organism type in Figure S15. The strongest study assessed for each organism type and test substance category combination is presented in Table 2.5.

Table 2.5 Arctic ecotoxicology studies with the greatest reliability scores for each organism type tested in the present study. Studies presented in bold are considered reliable for use in ecological risk assessment (overall reliability score ≥ 50%). Study number corresponds to the alphabetically-ordered list of assessed studies in Table S1.

Organism Type	Test Substance Category	Study No.	Study	Test Substance	Test Species	Marine or Freshwater	Endpoints	Overall Reliability Score (%)
Fish	Oil-related Contaminant	34	**Nahrgang et al., 2016**	Crude oil	*Boreogadus saida*	Marine	Cardiac activity and arrhythmia Hatching Malformations Mortality	87
	Inorganic Contaminant	30-33	Moore, 2016	Metals Metal salts Nitrogenous contaminants	*Coregonus clupeaformis* *Prosopium cylindraceum* *Salvelinus alpinus* *Salvelinus namaycush* *Thymallus arcticus*	Freshwater	Mortality	33*
	Organic Contaminant	40	**Palace et al., 2001**	PCB Congeners	*Thymallus arcticus*	Freshwater	EROD activity Osmolality Thyroid status Vitamin status	60*
Bird	Oil-related Contaminant	NA	NA	NA	NA	NA	NA	NA
	Inorganic Contaminant	2	**Braune et al., 2012**	Methylmercury	*Sterna paradisaea* *Uria lomvia*	Marine	Survival to 90% of development	80*
	Organic Contaminant	NA	NA	NA	NA	NA	NA	NA
Crustacean	Oil-related Contaminant	15a	**Gardiner et al., 2013**	Crude oil	*Calanus glacialis*	Marine	Mortality	80
	Inorganic Contaminant	39	**Øverjordet et al., 2014**	Mercury	*Calanus glacialis* *Calanus finmarchicus*	Marine	Gene expression (GST mRNA) Mortality	73

Organism Type	Test Substance Category	Study No.	Study	Test Substance	Test Species	Marine or Freshwater	Endpoints	Overall Reliability Score (%)
	Organic Contaminant	45b	Riebell and Percy, 1989	Phenol	*Mysis oculata*	Marine	Loss of equilibrium Lying on dorsal side Mortality No movement without prodding	12*
Algae	Oil-related Contaminant	4	**Camus et al., 2015**	Produced water	*Porosira glacialis*	Marine	Growth inhibition	67
	Inorganic Contaminant	NA	NA	NA	NA	NA	NA	NA
	Organic Contaminant	NA	NA	NA	NA	NA	NA	NA
Mollusc	Oil-related Contaminant	4	**Camus et al., 2015**	Produced water	*Mytulis edulis*	Marine	Development	67
	Inorganic Contaminant	48	Thyrring et al., 2015	Lead	*Mytulis edulis*	Marine	Fatty acid composition Mortality	30*
	Organic Contaminant	NA	NA	NA	NA	NA	NA	NA
Mammal	Oil-related Contaminant	8	Engelhardt, 1981	Crude oil	*Ursus maritimus*	Marine	Electrolyte changes Enzyme and metabolite changes Erythrocytic changes Hormone changes Leukocytic changes Nitrogenous compound changes	6*

Organism Type	Test Substance Category	Study No.	Study	Test Substance	Test Species	Marine or Freshwater	Endpoints	Overall Reliability Score (%)
	Inorganic Contaminant	14a and 14b	Frouin et al., 2012	Mercury	*Delphinapterus leucas*	Marine	Cell viability Intracellular thiol levels Lymphocyte proliferation Metallothionein induction	23*
	Organic Contaminant	6a and 6b	**Desforges et al., 2017**	Blubber-derived contaminant cocktail	Cetacean spp. *Orcinus orca* *Pagophilus groenlandicus* Seal spp. *Ursus maritimus* *Cystophore cristata*	Marine	Natural killer cell activity Cell viability Lymphocyte proliferation Phagocytosis	60
Microbe	Oil-related Contaminant	44	**Petersen et al., 2008**	Pyrene	Unidentified microbial consortium	Marine	Algal 14-C incorporation Chlorophyll-*a* content Phosphate flux Potential nitrification Potential oxygen consumption Silicate flux	60
	Inorganic Contaminant	NA	NA	NA	NA	NA	NA	NA
	Organic Contaminant	NA	NA	NA	NA	NA	NA	NA
Other Inv.	Oil-related Contaminant	9	Engelhardt, 1983	Crude oil	*Strongylocentrotus droebachiensis*	Marine	Curling of tube feet Loss of attachment ability Loss of cover Mortality Retraction of tube feed	30

Organism Type	Test Substance Category	Study No.	Study	Test Substance	Test Species	Marine or Freshwater	Endpoints	Overall Reliability Score (%)
							Spine droop Spine rigidity Stimulus response	
		42a	Percy and Mullin, 1975	Crude oil	*Halitholus cirratus*	Marine	Mortality	30
	Inorganic Contaminant	NA	NA	NA	NA	NA	NA	NA
	Organic Contaminant	NA	NA	NA	NA	NA	NA	NA

*The only study for this category.

2.3.7 Relevance Scores

Relevance scores for all experimental combinations assessed in the present study are presented by organism type in Table 2.6. The mean relevance score for fish tests (with the greatest mean relevance score; 5.3 ± 1.4) and mammalian tests (with the least mean relevance score; 2.0 ± 0) were both significantly different from all other organism groups and each other ($p < 0.05$).

Table 2.6 Count of relevance scores for all reported responses assessed in this study (i.e., test substance/test species/endpoint combinations, *n* = 596), except birds and microbes, by organism type.

Relevance Score	Crustacean (*n* = 276)	Fish (*n* = 173)	Algae (*n* = 17)	Mammal (*n* = 35)	Mollusc (*n* = 23)	Other Inv. (*n* = 15)
0	0	0	0	0	0	0
1	10	2	0	0	0	0
2	2	18	0	35	3	0
3	60	0	0	0	9	11
4	33	17	15	0	1	0
5	20	4	0	0	1	0
6	151	132	2	0	9	4

Crustacean tests achieved the second-highest mean relevance score (4.8 ± 1.5), followed by molluscs (4.2 ± 1.6), algae (4.2 ± 0.7), and other invertebrates (3.8 ± 1.4). No statistically significant differences were found between any other organism groups other than those described above. When assessed by test substance type, differences between mean relevance scores for inorganic contaminants (5.5 ± 1.2), oil-related contaminants (4.7 ± 1.6), and organic contaminants (2.3 ± 1.0) were all statistically significant from one another (Figure S16). Figure 2.6 provides the relationship between overall and relevance scores for the entire dataset as a whole (*n* = 596 experimental combinations). The points in the lower left quadrant indicate both low overall and relevance scores (20% of experimental combinations). The upper left quadrant contains studies that used highly relevant endpoints but did not score well overall (59% of experimental combinations). The lower right quadrant indicates high overall reliability

scores but with minimal relevance for use in risk assessment (7% of experimental combinations). The upper right quadrant includes studies that are considered highly relevant, achieved a score of 1 in all the critical categories outlined above, and were generally well-designed, conducted, and reported (14%). This relationship is broken down further by organism group in Figure S17.

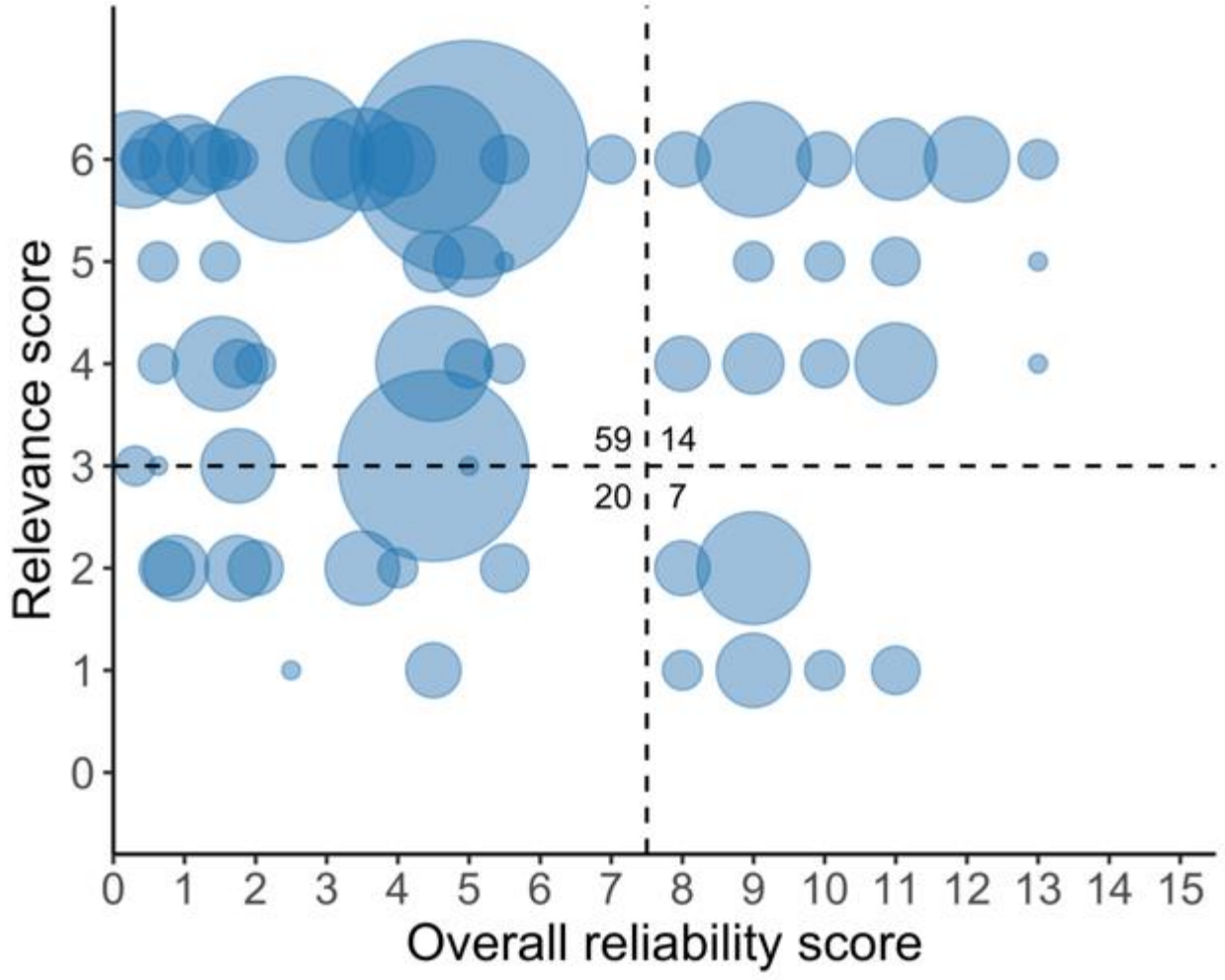

Figure 2.6 Strength of methods and relevance scores for all reported responses in this study (*n* = 596). Size of circle corresponds to number of studies. Numerical values in each quadrant refer to percentage of data points with a relevance score between 0 and 3 (inclusive) and an overall reliability score of < 50% (bottom left quadrant) or > 50% (bottom right quadrant), and relevance score between 4 and 6 (inclusive) and an overall reliability score of < 50% (top left quadrant) or > 50% (top right quadrant).

2.4 Discussion

This review supports the data reliability concerns that have been raised by the ecotoxicological community in recent years (e.g., Ågerstrand et al., 2013; Breitholtz et al., 2006; Durda and Preziosi, 2000; Harris et al., 2014) as it pertains to Arctic species toxicity testing. A specific sensitivity comparison between Arctic and temperate species was not a focus of this review; however, it is a question that needs to be answered for environmental risk assessors conducting ERAs. Overall, relatively few Arctic species toxicity tests assessed herein achieved an overall reliability score deemed minimally acceptable for use in risk assessment (46 of 253 unique toxicity tests; 18%) without significant uncertainty; therefore, there are limited data available to use in a sensitivity comparison. Of those that achieved acceptable scores, there are many important groups of test species, test substances, and endpoints that are highly underrepresented relative to their potential for exposure in the Arctic, which are discussed below. The results of the present study have highlighted gaps in collected data and clearly demonstrated the need for more methodologically robust studies and have provided a set of collected data that could be used to address these relative sensitivity questions in the future. Additionally, this study presented an opportunity to further refine the scoring rubric, discussed below.

2.4.1 General Recommendations for Improving the Conduct of Future Toxicity Studies

Several methodological deficiencies across studies came to light during the scoring process. Firstly, appropriate controls should always be included with every test, regardless of prior evidence obtained that supports the decision to exclude them. A common approach in the studies assessed here was to indicate that a solvent was used to dissolve the test substance, but that no solvent control was included due to evidence in previous studies that such concentrations did not significantly impact the organism. Errors in solution preparation, solvent stock preparation, or increased sensitivity of a particular group of test organisms cannot be definitively ascertained without the inclusion of these controls. Ultimately, each test compound may have a unique

interaction with a solvent control that needs to be accounted for in the test design (Christou et al., 2020).

Secondly, reporting of environmental conditions needs significant improvement to ensure that tests can be repeated and results reproduced. For example, studies often reported the target temperature or lighting of exposure chambers without reporting actual measured values. Issues can arise with growth chambers, from temperature and lighting swings and inconsistencies spatially, to faulty light bulbs that stop working midway through exposure, to power failures that may not be recognized if no monitoring is conducted. In some cases, monitoring may have been conducted but results not reported in the publication itself. Although this information may have been retrievable through contacting the authors to confirm whether or not certain monitoring was done or whether measured values were within the pre-defined range, it is unrealistic and impractical to do this when it could easily be included in the original paper. Just as it is critical to analytically confirm exposure concentrations to have confidence in the results, the user of the data will not have confidence that test conditions were within the pre-defined range for the duration of exposure if target conditions are not reported as well as measured.

Thirdly, low exposure volumes can lead to an inability to analytically confirm exposure concentrations (e.g., algae and copepod tests in Camus et al., 2015). The approach taken in this study was to estimate measured concentrations by using the average ratio of nominal to measured concentrations that were reported for other tests in the same study. This likely provides a more accurate estimation of the actual exposure than nominal concentrations alone; however, it does not give risk assessors full confidence in the results as if the individual unit concentrations were reported. Including larger volume surrogate vessels that are exposed for the same duration and under the same conditions as EUs can be an option for these scenarios.

Finally, and perhaps most importantly, it is imperative to stress that not all of the studies assessed herein claimed to have been conducted with the objective of toxicity data for use in ERA. Understanding the biological traits of Arctic species through a variety of study designs is equally as critical to achieving the overall goal of environmental sustainability as is generating toxicity data; however, the aim of this study

was to identify studies exclusively for the latter purpose. A poor score achieved through the application of this rubric does not necessarily mean that the study was not well-conducted or of high value to science; rather, it implies that the study does not necessarily align with the sole function outlined *a priori* in this review. For example, criterion 5 indicates that a minimum of three concentrations must be tested. This is the minimum number of concentrations required to plot a concentration-response curve to characterize the range of toxicity to an organism. This is useful for ERA, but may not be required to produce scientifically valuable results. If a study were assigned a score of 0 for this criterion because only a single concentration was tested to assess the mode of action of a compound, the data may not carry as much weight in the risk assessment sphere though the results are important to science. Understanding the mechanism of action of various contaminants upon exposure to Arctic species in Arctic conditions or generating data that will assist in our knowledge of biomarkers of exposure in the field is extremely relevant. These studies contribute to the foundation of knowledge that is required to advance the field of ecotoxicology and ERA and would likely render toxicity data less useful if it were not gathered. We recommend that these type of data continue to be collected, but it would be beneficial if studies were more explicit in their objectives as to the purpose of the data collection so that data collected for the purpose of informing risk assessment could be easily identified.

2.4.2 Specific Recommendations to Advance Arctic Toxicity Testing

This research has highlighted several knowledge gaps and deficiencies in reporting that should be addressed for future published work. The recommendations below should be taken into consideration by anyone planning, performing, or reviewing ecotoxicological data including researchers, funding agencies, and risk assessors.

2.4.2.1 Standardization of Arctic-specific test methods

Knowledge Gap #1: Lack of standard toxicity and recovery test methods for Arctic species.

Recommendation #1: Strive toward developing standard laboratory tests (whole organism and *in vitro*), and in the absence of standard tests provide a thorough description of what was modified from standard methods.

Prior to discussing the dataset as a whole, it is important to note that standard toxicity test methods designed specifically for Arctic aquatic species do not currently exist in the literature. There have been techniques introduced by researchers for modelling Arctic-specific toxicity data that take into account environmental, geophysical, and physiological factors specific to Arctic regions that partially bridge this gap; however, they are novel in their application (e.g., Fahd et al., 2019). Many of the studies scored in this review followed standard test methods as closely as possible, but with modifications to environmental parameters, test durations, or validity criteria as necessary to accommodate the requirements of Arctic test organisms (e.g., Camus et al., 2015; Faksness et al., 2011; Hansen et al., 2012; Riebell and Percy, 1989). However, several studies reported that standard methods were followed but with "modifications", without detailing what was modified from the original test conditions. This renders the experiment impossible to reproduce or score effectively. Researchers should at least report a brief summary of the standard method conditions instead of referencing a separate document with or without modifications.

Developing and validating standard *in vivo* tests for Arctic regions is necessary as data based on modelling or collected via *in vitro* methods need to be validated by *in vivo* studies to have confidence in the results. Additionally, there are many questions that cannot be answered with tests that do not expose the whole organism, such as behaviour and physiological interactions within the organism. The current regulatory infrastructure relies on whole-organism based endpoints (e.g., mortality) to determine hazardous concentrations in the environment and thus these data are enormously helpful to ERA. In short, both *in vivo* and *in vitro* tests are important to conduct, each offering a range of benefits and drawbacks that the other does not.

As it pertains specifically to *in vitro* tests, emphasis should be placed on developing methods where the sensitivity, specificity, and exposure conditions of the assays are clearly understood and defined (Rehberger et al., 2018). The purpose of

these tests, including how the data can be extrapolated to the ecosystem level, should be plainly expressed by researchers to facilitate understanding by environmental risk assessors regarding how the data can be used in the context of ERA.

Finally, understanding potential latent effects and/or the capacity of test organisms to recover post-exposure is a critical component in extrapolating ecosystem-level effects from lab-based assays; however, only a quarter of tests assessed herein included post-exposure monitoring (Beketov and Liess, 2009; Pechenik, 2006). This may be even more important for Arctic test organisms as a variety of biotic and abiotic factors specific to this ecosystem result in significant uncertainty regarding time to effect (Chapman and Riddle, 2005; Olsen et al., 2011). For example, longer toxicity response times for polar invertebrates relative to temperate ones have been observed for metal exposures. It has been hypothesized that this can be partially attributed to the temperature-driven low metabolic rate of Arctic species as well as the longer development time of polar organisms relative to their temperate and tropical counterparts. Additionally, sea-ice cover is known to dampen wind-driven mixing of the water column, which increases the probability of long exposure times for under-ice organisms (Chapman and Riddle, 2005; Cota et al., 1991).

2.4.2.2 Test species

Knowledge Gap #2: Lack of toxicity data for Arctic-focused risk assessments in general, especially for some ecologically important groups.

Recommendation #2: Generate toxicity data using organisms that are likely the most sensitive first (i.e., algae). Emphasize the development and refinement of models and surrogates (e.g., temperate species) using these data and confirm protection with ongoing toxicity testing.

The diversity of Arctic species that have been tested in the studies assessed herein is not directly proportionate to the diversity of species that live in this region, though there are some parallels. CAFF (2013) reported that of the ~7,630 Arctic marine species (mammals, seabirds, fish, invertebrates, and algae) reported in their 2013 Arctic

Biodiversity Assessment, 65.5% were invertebrates. In the present study the distribution of tested species aligned with these data as ~67% of marine species tested were invertebrates (*n* = 43 of 67 marine species). In contrast, of all fish species reported in the Arctic, ~34% were freshwater and diadromous while the remaining 66% were marine. In the present study, 42% of tested fish species were from freshwater or diadromous environments and 58% were marine. This was largely driven by a study that focused on the impacts of mining contaminants on freshwater fish (Moore, 2016). Marine algae were underrepresented in the assessed literature relative to environmental populations, with only 6 of 146 marine tests being conducted on these organisms (4% of marine tests) while marine algae are the second-most diverse species group in the Arctic based on the same report (~30% of all marine species). The preference to study some species groups over others (i.e., taxonomic chauvinism) has been observed in ecotoxicology for pesticides and may occur for a variety of reasons (Prosser et al., 2021). Within Arctic-specific toxicology, the lack of data using algae as test species may be due to the limited availability of ecologically important species in commercial culture collections.

Tests conducted with fish were plentiful; however, the majority were conducted by a single group (Moore, 2016) and thus there is limited reproducibility of these results. Scores for fish were also driven by this single study and, as mentioned above, could potentially have vastly different scores upon secondary review. Further, only 17% of fish tests (*n* = 21 of 126) were conducted using marine species, an area of research where more testing would be beneficial due to the abundance and ecological importance of marine fish in the Arctic, their role as a country food for Indigenous people of the North, and the anticipated increase in Arctic shipping that will increase their likelihood of exposure to contaminants (Arctic Council, 2020, Darnis et al., 2012).

Relatively few data points were collected from mammals, molluscs, other invertebrates, microbes, and birds. Of these, we believe the priority should be placed on collecting algal data. With a relatively low number of data points derived from algae in this review (*n* = 25), they also achieved the lowest scores in two of the three criteria groups: Group A (test substance) and Group C (test design, statistics, and results). For

algae, it is a case of both lack of data, and issues with reliability that need to be addressed.

The reasoning for prioritizing on algal testing is two-fold. The first is their relative sensitivity and availability. Algae have demonstrated sensitivity to many substances (Blaise and Vasseur, 2005) and are frequently used by environmental risk assessors to help define regulatory thresholds of contaminants in temperate ecosystems. They can also be cultured year-round in the laboratory with minimal cost and infrastructure. Second, the importance of ice algae in Arctic ecosystems cannot be overstated. For example, diatoms such as *N. frigida* and *Melosira arctica* are drivers of production in the Arctic Ocean, forming a seasonally critical energy pulse for the food web when they bloom in spring. (Fernández-Méndez et al., 2015; Gosselin et al., 1997; Michel et al., 1996; Søreide et al., 2010). These resources are extremely important in sustaining the structure and function of Arctic ecosystems and they are severely underrepresented in Arctic ecotoxicology.

Another option to address the gap in Arctic species' toxicity testing would be to place emphasis on modelling and other alternative methods to estimate toxicity. For example, the Interspecies Correlation Estimation model (WebICE; Raimondo et al., 2015) predicts the toxicity of contaminants to untested organisms based on surrogate species and has great potential to be refined further to increase its reliability (Bejarano et al., 2017a; Bejarano, 2019; Bejarano and Wheeler, 2020). Additionally, SSDs and hazard concentrations (HCs) for Arctic and non-Arctic species have been compared in the literature and have been reported to be protective of Arctic organisms for select compounds (e.g., hydrocarbons; Bejarano et al., 2017b). Further generation of high-quality toxicity data for other compound groups would contribute to these models thus allowing risk assessors more confidence in extrapolating results from such data.

Though there are limited data, some comparisons of Arctic versus temperate species sensitivity have been made in the literature. For example, Olsen et al. (2011) compared sensitivities of Arctic and temperate organisms to 2-methylnapthalene and found that they were similarly sensitive. Similarly, hazardous concentrations of artificial produced water to Arctic and temperate fish have been found to be comparable (Camus et al., 2015). Hansen et al. (2014) found that neither Arctic nor temperate organisms

were consistently the most sensitive to oil spill response chemicals. These comparisons that have already been generated in the literature demonstrate no clear, repeatable evidence that one group is more, less, or similarly sensitive than the other, underlining the importance of generating high quality data to facilitate these comparisons effectively.

2.4.2.3 Test substances

Knowledge Gap #3: Lack of diversity in compounds tested on Arctic organisms.

Recommendation #3: Design repeatable and reliable experiments using standard reference toxicants so that results can be compared across space and time.

Knowledge Gap #4: Inconsistencies in oil spill related research has led to a lack of comparability between studies, Arctic or otherwise.

Recommendation #4: Focus research on developing, validating, and refining laboratory methods and predictive effects models.

A wider variety of test substances that are representative of contaminants found in Arctic environments should be incorporated into future research. The vast majority of our knowledge thus far is related to oil spills and oil spill remediation methods. There have been no studies assessing the effect of multiple stressors on Arctic organisms, nor other emerging contaminants besides a single study on microplastics (Rodríguez-Torres et al., 2020). For example, the impact of contaminants found in wastewater have not been determined via first-tier, laboratory toxicity tests (e.g., PPCPs), nor have POPs (e.g., pesticides and industrial chemicals). It is unreasonable to suggest that all compounds that fit within these categories should be tested, but several contaminants representative of these groups for which there are background data in the literature would suffice to increase our understanding of how organisms may be impacted by

exposure. Further, as regulations for Arctic activities change on a political, economic, and social scale, new contaminants of concern will continue to emerge that will require action by the ecotoxicological community. Data has been collected from field monitoring exercises (e.g., tissue concentrations from naturally exposed marine mammals compared to biomarkers of exposure within the organism; see Table S1); however, this does not provide the data that risk assessors require to inform regulatory decision making under the current framework.

More than half of all collected data uses oil-related contaminants alone. The vast estimates of untapped hydrocarbons in Arctic regions and increased shipping traffic over the last decade provide sound reasons to study the impacts of potential future oil spills in the marine environment (United States Department of the Interior, 2008; Ostreng et al., 2013). The potential for Arctic organisms to be exposed to both local and foreign oils as a result of Arctic hydrocarbon exploration and fuel transportation exists and was well-captured in the studies assessed in this review. As of January 2020, the International Maritime Organization (IMO) has implemented a new limit on fuel oil sulphur content to reduce sulphur oxide emissions from ships, which introduces a new potential contaminant threat to Arctic organisms and presents an opportunity for identifying additional test substances to use in toxicity testing (International Maritime Organization, 2019). This underscores the importance of continuously adapting toxicity tests to ensure the data being collected are needed to answer current questions being asked by environmental risk assessors as regulations change.

Though the question of oil-related contaminant toxicity has been partially answered with the tests assessed in this review, inconsistent methodologies and reporting in oil spill research have limited our ability to compare and use the data generated by these experiments. This issue has been raised by researchers in the past and some have provided guidance for improving consistency in the future (e.g., Bejarano et al., 2019; Redman and Parkerton, 2015). Deviations from the standard methods are common, often with limited or no description of what said deviations were as described above. Although protocols do exist to address some of these difficulties (e.g., Chemical Response to Oil Spills Ecological Effects Research Forum [CROSERF] method; Aurand and Coelho, 2005), the complex nature of oil as a test substance

continues to provide challenges in exposure media preparation, characterization, and reporting. The myriad of compounds that make up crude oil render it problematic to characterize the toxicity of every one, and a wide variety of environmental, physical, and biological factors can have extreme influence on how the compounds interact with one another and with the test organism (Wang et al., 2021). Predictive models have become more popular in the last decade (Carroll and Smit, 2011; Hansen et al., 2019; McGrath et al., 2018) and are a promising approach to addressing this issue.

There were many test substances that were used across multiple studies on the same or similar organisms (e.g., crude oil tested on crustaceans in Busdosh and Atlas, 1977, Gardiner et al., 2013, and Hansen et al., 2009; produced water tested on fish in Camus et al., 2015 and Geraudie et al., 2014); however, most compounds where data exist for similar species across multiple studies are difficult or impossible to compare as the test substances are complex mixtures that differ depending on their source. Pyrene is one exception; this compound was tested on crustaceans in six studies (Grenvald et al., 2013, Hjorth and Nielsen, 2011, Jensen et al., 2008, Nørregaard et al., 2014, Toxværd et al., 2018, and Toxværd et al., 2019) but with major differences in test organism and endpoint parameters (i.e., life stages, feeding strategies (starved/feeding), ECx values calculated at various time points that do not temporally align) that again render comparison difficult. This variability demonstrates that risk assessors would benefit from the generation of a bank of Arctic-specific toxicity data that uses standard reference toxicants so that tests can be shown to be repeatable. ECCC (1990) has outlined several criteria to determine the acceptability of various reference toxicants, including the ability to detect abnormal organisms, their solubility and stability over time, their ease of analysis, and whether they have an established toxicity database. Not only does the inclusion of reference toxicants in toxicity testing facilitate inter- and intra-laboratory comparisons of results through space and time, but also enables laboratories to determine the precision of their toxicity tests (ECCC, 1990). This will also ensure that results can be confidently compared to understand where the sensitivity of Arctic organisms lie relative to standard, temperate species when they are employed as surrogates.

2.4.2.4 Endpoints

Knowledge Gap #5: Lack of highly relevant sublethal and chronic endpoints assessed in Arctic toxicity testing.

Recommendation #5: Prioritize the incorporation of sublethal and chronic endpoints into future Arctic toxicity test protocols.

More than half of all endpoints reviewed in this paper were mortality. Although these data are necessary and useful to collect, sublethal and chronic responses are severely underrepresented. Understanding the threshold of concentrations where organisms will die (e.g., LC50 values) is critical for calculating risk to the ecosystem as a whole; however, it is important to prevent contaminant concentrations from reaching these levels in the environment. Mortality may be the most highly relevant endpoint identified in this rubric for risk assessments, but chronic exposures provide information that allow risk assessors to remain as conservative as possible in their recommendations to regulatory decision-makers. The US EPA (1994) has highlighted the ecological significance of laboratory-derived sublethal effect data when extrapolated to field conditions. Growth, developmental, and reproductive changes from a population's norm have the capability to drastically alter a population's biomass and age structure. Examples of sublethal endpoints that have been recommended in the literature for future study range from the organism level (e.g., avoidance, courtship, migratory, and rearing behaviour) to the community, ecosystem, and landscape level (e.g., nutrient retention, decomposition rate, resilience and recovery; US EPA, 2016). The only organism type from studies assessed in this review that were primarily studied for non-lethality endpoints was the other invertebrate category, for which behavioural endpoints comprised most of the data; therefore, algae, birds, crustaceans, fish, mammals, microbes, and molluscs are lacking these data. Where some work has already been started on developing methods to test sublethal endpoints with Arctic species (e.g., reproductive tests with the copepod *Calanus glacialis* [Jensen et al., 2008] and the fish *Mallotus villosus* [Frantzen et al., 2012] or behavioural tests with the

mollusc *Mya truncata* and the echinoderm *Strongylocentrotus droebachiensis* [Engelhardt, 1983]), these species would be excellent candidates for further developing and refining said methods in order to produce high quality, reliable data. A summary of the key knowledge gaps and recommendations can be found in Table S5.

2.4.3 Critiquing and Modifications to the Rubric

The scoring rubric used in this study was modified from that described in Hanson et al., (2019) for the herbicide atrazine and underwent modifications based on the specific Arctic ecotoxicological data that had been collected. This ensured that the spreadsheet was appropriately designed for Arctic-specific data and included space for a range of exposure scenarios. That said, there were still studies that did not fit the spreadsheet fully and could not be objectively scored with total confidence according to the criteria set out in advance (e.g., *in vitro* studies, abbreviated abstracts that purposefully did not include all relevant methodologies but were not published in full elsewhere, studies that did not include three replicates but were still statistically sound). These studies would benefit from a secondary or consensus review where expert judgment could be used to determine scores more accurately. Though potential issues and modifications to the spreadsheet for the future are noted here, a full secondary review is outside of the scope of this paper.

Firstly, a minimum of three replicates per exposure is necessary to meet criterion 11. Achieving a score of 0 in this criterion, as it is considered critical to the study, results in the application of a multiplication factor of 0.5 to the overall reliability score. Although appropriate replication is in fact critical to the study, three replicates are not an absolute requirement, depending on study design (e.g., studies using a linear regression approach as per the ECCC test method [ECCC, 2000]), to be deemed appropriate replication. For example, the fact that relatively few experiments met this criteria was largely driven by a single study that produced 104 of the 126 fish tests (Moore, 2016). This study followed the ECCC test method that requires twenty fish to be placed in a single bucket with no replicate buckets for each concentration, a method that has been deemed statistically appropriate and acceptable due to the statistical power of the test. Due to the strictness of the criterion, this study resulted in an overall reliability score of

5/15 after the multiplication factor was applied for scoring 0 in criterion 11, when upon secondary review it may have been scored as 10/15. If the criterion indicated that a minimum of 3 replicates per exposure were required OR that justification of an otherwise appropriate statistical approach could still warrant a score of 1, this study would have been considered as reliable for use in risk assessment (i.e., overall reliability score > 50%).

Secondly, the rubric includes three separate criteria related to chemical analyses of exposure media: any analytical confirmation and reporting of tested concentrations (e.g., stock solutions), initial concentration in individual test units reported, and final concentrations in individual test units reported. While this is useful and sometimes necessary information to report, the necessity should be determined based on the test compound with an opportunity to adjust the score if required. For example, many crude oils contain compounds that can volatize relatively quickly such as toluene and benzene, thus exposure media using oil as a test substance can change significantly over the duration of the test and these toxicity-inducing compounds may no longer be present (Rajabi et al., 2020). In this scenario, it is necessary to chemically analyze the media at both test initiation and termination to have confidence in the exposure. However, metals such as copper and zinc have been used extensively in the ecotoxicological literature and have been recommended as reference toxicants by ECCC (1990) due to their demonstrated ability to maintain consistent concentrations over long periods of time, among other criteria. In this scenario, confirming concentrations at the start of exposure is likely sufficient due to the reasons outlined above. In future assessments of data reliability, a table could be created that outlines appropriate chemical analyses and reporting requirements for different classes of compounds based on best available scientific practice and need.

Thirdly, the rubric did not lend itself well to the assessment and scoring of studies conducted *in vitro*. The purpose and approach of these tests can be significantly different than a whole-organism toxicity test, resulting in fundamental differences in what criteria may be considered critical to the reliability of the study. For example, information required to be reported for whole-organism toxicity tests may include environmental parameters such as temperature, light intensity, and photoperiod,

husbandry practices including media renewal and feeding strategies, and control performance criteria. Most of these requirements are not applicable for tests conducted on cell lines as the test organism husbandry element has been removed, so objectively scoring becomes difficult. Additionally, *in vitro* tests typically compare treatments to unexposed cell lines used as controls (e.g., Desforges et al., 2017). This allows results to be expressed as percent of control values; however, there should be some sort of control criteria for *in vitro* tests that allows researchers to compare both treated and control cell lines to a predefined value. For this reason, these tests did not achieve a score of 1 in criterion 15 (control values reported and performance criteria met) even though the control performance may have been acceptable for this type of test. In future applications of this scoring rubric, *in vivo* and *in vitro* studies should be assessed separately according to criteria that have been predefined that are specific to each group. Several tools exist for evaluating data collected via *in vitro* methods, including ToxRTool and SciRAP, could be useful in informing this aspect of the rubric in the future (Roth et al., 2021; Schneider et al., 2009).

Lastly, it would be beneficial to develop more distinct criteria in future versions of this rubric or for scoring in general. For example, criterion 12 (statistical methods described) indicates that the appropriate test must be reported and appropriate controls employed. This criterion could be split into two separate criteria so that each of these issues could be addressed separately. As it stands, if the test employed was not described appropriately OR if appropriate controls are not included, the study does not achieve a score of 1 in that criterion when it may have actually met one or the other. Implementing these changes would give risk assessors a more accurate picture of the experiment with all relevant details easily identifiable from the rubric. Ultimately there is a balance that needs to be struck between interrogating all aspects of a study with expediency and a focus on those that speak most directly to methodological strength and reliability for risk assessment.

Despite the uncertainties resulting from the application of the scoring rubric discussed here, the overall objective of the exercise was met. The transparent nature of defining criteria *a priori* meant that once the method was defined, it was followed for all studies to ensure consistency. This resulted in some unavoidable rigidity and there

could be changes to several scores upon the implementation of suggestions described above; however, the original scores and rubric have been included to assist other researchers in assessing and modifying, as needed, the reliability of studies for use in ERA as they deem fit. The process of developing a method such as this is iterative and can only be improved after working with it and acknowledging uncertainties that may arise.

2.5 Conclusion

We found that the application of the scoring rubric identified methodologically strong, reliable Arctic ecotoxicology studies that could be prioritized for use in ERA and regulatory guidelines. The exercise determined that the vast majority of studies have significant weakness and uncertainties, and should likely be recognized as such when aiming to inform regulatory decisions. This highlights the need for risk assessors and ecotoxicology researchers in general to critically evaluate the reliability of data before any decision-making occurs. To save time, energy, and resources, we encourage researchers to consult a rubric such as the one presented in this study before designing their own experiments to ensure that minimum experimental and reporting requirements for high quality data are met, as well as need for the toxicity test itself. Performing a test that is poorer in design and methodology than past work does not move environmental protection forward, but rather undermines the effort, and should be avoided.

2.6 References

Ågerstrand, M., Küster, A., Bachmann, J., Breitholtz, M., Ebert, I., Rechenberg, B., Rudén, C. (2011a). Reporting and evaluation criteria as means toward a transparent use of ecotoxicity data for environmental risk assessment of pharmaceuticals. *Environmental Pollution, 159*(10), 2487-2492.

Ågerstrand, M., Breitholtz, M., Rudén, C. (2011b). Comparison of four different methods for reliability evaluation of ecotoxicity data – A case study of non-standard test data used in environmental risk assessments of pharmaceutical substances. *Environmental Science Europe, 23*, 17.

Ågerstrand, M., Edvardsson, L., Rudén, C. (2013). Bad reporting or bad science? Systematic data evaluation as a means to improve the use of peer-reviewed studies in risk assessments of chemicals. *Human and Ecological Risk Assessment: An International Journal, 20*(6), 1427-1445.

Arctic Council. (2020). *First Arctic shipping status report from PAME working group highlights increase in Arctic shipping traffic.* https://arctic-council.org/en/news/first-arctic-shipping-status-report-increase-shipping-traffic/.

Arctic Monitoring and Assessment Programme (AMAP). (2017). *Adaptation actions for a changing Arctic: Perspectives from the Bering-Chukchi-Beaufort region.* Arctic Monitoring and Assessment Program (AMAP). Retrieved from https://www.amap.no/documents/download/2993/inline.

Aulanier, F., Simard, Y., Roy, N., Gervaise, C., and Bandet, M. (2017). Effects of shipping on marine acoustic habitats in Canadian Artic estimated via probabilistic modelling and mapping. *Marine Pollution Bulletin, 125*(1-2), 115-131.

Aune, M., Aschan, M. M., Greenacre, M., Dolgov, A. V., Fossheim, M., Primicerio, R. (2018). Functional roles and redundancy of demersal Barents Sea fish: Ecological implications of environmental change. *PloS One, 13*(11), e0207451.

Aurand, D. and Coelho, G. (Editors). (2005). C*ooperative aquatic toxicity testing of dispersed oil and the "Chemical Response to Oil Spills: Ecological Effects*

Research Forum (CROSERF)." Ecosystem Management and Associates, Inc. Lusby, MD. Technical Report 07-03.

Bejarano, A. (2019). Further development and refinement of interspecies correlation estimation models for current-use dispersants. *Environmental Toxicology and Chemistry, 38*(8), 1682-1691.

Bejarano, A. C. and Wheeler, J. R. (2020). Scientific basis for expanding the use of interspecies correlation estimation models. *Integrated Environmental Assessment and Management, 16*(4), 528-530.

Bejarano, A. C., Coelho, G. M., Mitchelmore, C. L., Wetzel, D. L. (2019. November 3-7). *A path forward for improving aquatic toxicity testing of oil* [conference proceedings]. Annual Meeting of the Society of Environmental Toxicology and Chemistry, Toronto, Ontario, Canada.

Bejarano, A. C., Gardiner, W. W., Barron, M. G., and Word, J. Q. (2017b). Relative sensitivity of Arctic species to physically and chemically dispersed oil determined from three hydrocarbon measures of aquatic toxicity. *Marine Pollution Bulletin, 122*(1-2), 316-322.

Bejarano, A., Raimondo, S., Barron, M. G. (2017a). Framework for optimizing selection of interspecies correlation estimation models to address species diversity and toxicity gaps in an aquatic database. *Environmental Science and Technology, 51*(14), 8158-8165.

Beketov, M. A. and Liess, M. (2009). Acute and delayed effects of the neonicotinoid insecticide thiacloprid on seven freshwater arthropods. *Environmental Toxicology and Chemistry, 27*(2), 461-470.

Blaise, C. and Vasseur, P. (2005). Algal Microplate Toxicity Test. In C. Blaise and J.-F. Ferard (Eds.), *Small-Scale Freshwater Toxicity Investigations* (pp. 137-179). Springer, Dordrecht.

Borgert, C. J., Becker, R. A., Carlton, B. D., Hanson, M., Kwiatkowski, P. L., Marty, M. S., McCarty, L. S., Quill, T. F., Solomon, K., van der Kraak, G., Witorsch, R. J., Yi, K. D. (2016). Does GLP enhance the quality of toxicological evidence for regulatory decisions? *Toxicological Sciences, 151*(2), 206-213.

Braune, B. M., Scheuhammer, A. M., Crump, D., Jones, S., Porter, E., and Bond, D. (2012). Toxicity of methylmercury injected into eggs of thick-billed murres and Arctic terns. *Ecotoxicology, 21,* 2143-2152.

Breitholtz, M., Rudén, C., Hansson, S. O., Bengtsson, B.-E. (2006). Ten challenges for improved ecotoxicological testing in environmental risk assessment. *Ecotoxicology and Environmental Safety, 63*, 324-335.

Brocke, J., Simons, A., Niehaves, B, Niehaves, B., Reimer, K. (2009, June 8-10). *Reconstructing the giant: On the importance of rigour in documenting the literature search process* [conference proceedings]. European Conference on Information Systems (ECIS), Verona, Italy.

Brown, R. G. B. (1980). Seabirds as Marine Animals. In J. Burger, B. L. Olla, and H. E. Winn (Eds.), *Behaviour of Marine Animals* (pp. 1-39). Springer, Boston, MA.

Burden, N., Benstead, R., Benyon, K., Clook, M., Green, C., Handley, J., Harper N., Maynard, S. K., Mead, C., Pearson, A., Ryder, J., Sheahan, D., van Egmond, R., Wheeler, J. R., Hutchinson, T. H. (2020). Key opportunities to replace, reduce, and refine regulatory fish acute toxicity tests. *Environmental Toxicology and Chemistry, 39*(10), 2076-2089.

Busdosh, M. and Atlas, R. M. (1977). Toxicity of oil slicks to Arctic amphipods. *Arctic, 30*(2), 85-92.

Campbell, L. M., Norstrom, R. J., Hobson, K. A., Muir, D. C. G., Backus, S., Fisk, A. T. (2005). Mercury and other trace elements in a pelagic Arctic marine food web (Northwater Polynya, Baffin Bay). *Science of the Total Environment, 351-352,* 247-263.

Camus, L., Brooks, S., Geraudie, P., Hjorth, M., Nahrgang, J., Olsen, G. H., Smit, M. G. D. (2015). Comparison of produced water toxicity to Arctic and temperate species. *Ecotoxicology and Environmental Safety, 113,* 248-258.

Carroll, J. and Smit, M. (2011, February 22-24). *An integrated modeling framework for decision support in ecosystem management: Case study Lofoten/Barents Sea* [conference proceedings]. European Health, Safety, and Environmental Conference in Oil and Gas Exploration and Production, Vienna, Austria.

Chapman, P. M., and Riddle, M. J. (2003). Missing and needed: Polar marine ecotoxicology. *Marine Pollution Bulletin, 46*(8), 927–928.

Chapman, P. M., and Riddle, M. J. (2005). Polar marine toxicology – Future research needs. *Marine Pollution Bulletin*, *50*(9), 905–908.

Christou, M., Kavaliauskis, A., Ropstad, E., Fraser, T. W. K. (2020). DMSO affects larval zebrafish (*Danio rerio*) behaviour, with additive and interaction effects when combined with positive controls. *Science of the Total Environment, 709*, 134490.

Clarivate Analytics. (2021). Journal Citation Reports Science Edition.

Conservation of Arctic Flora and Fauna. (2013). Arctic Biodiversity Assessment: Status and trends in Arctic biodiversity. Conservation of Arctic Flora and Fauna, Akureyri.

Cota, G. F., Legendre, L., Gosselin, M., Ingram, R. G. (1991). Ecology of bottom ice algae: I. Environmental controls and variability. *Journal of Marine Systems, 2*, 257-277.

Darnis, G., Robert, D., Pomerleau, C., Link, H., Archambault, P., Nelson, R. J., Geoffroy, M., Tremblay, J.-E., Lovejoy, C., Ferguson, S. H., Hunt, B. P. V., Fortier, L. (2012). Current state and trends in Canadian Arctic marine ecosystems: II. Heterotrophic food web, pelagic-benthic coupling, and biodiversity. *Climatic Change, 115*(1), 179-205.

Desforges, J.-P., Levin, M., Jasperse, L., De Guise, S., Eulaers, I., Letcher, R. J., Acquarone, M., Nordoy, E., Folkow, L. P., Jensen, T. H., Grondahl, C., Bertelsen, M. F., Leger, J. S., Almunia, J., Sonne, C., Dietz, R. (2017). Effects of polar bear and killer whale derived contaminant cocktails on marine mammal immunity. *Environmental Science and Technology, 51,* 11431-11439.

Durda, J. L. and Preziosi, D. V. (2000). Data quality evaluation of toxicological studies used to derive ecotoxicological benchmarks. *Human and Ecological Risk Assessment, 6*(5), 747-765.

Engelhardt, R. (1983, June 14-16). *Behavioural responses of benthic invertebrates exposed to dispersed crude oil* [conference proceedings]. Arctic Marine Oil Spill Program Technical Seminar, Edmonton, Alberta, Canada.

Environment and Climate Change Canada (ECCC). (1990). *Guidance document on control of toxicity test precision using reference toxicants.* Report EPS 1/RM/12.

Environment and Climate Change Canada (ECCC). (2000). *Biological test method: Acute lethality of effluents to rainbow trout.* Report EPS 1/RM/13.

Fahd, F., Veitch, B., Khan, F. (2019). Arctic marine fish 'biotransformation toxicity' model for ecological risk assessment. *Marine Pollution Bulletin, 142,* 408-418.

Faksness, L.-G., Borseth, J. F., Baussant, T., Hansen, B. H., Altin, D., Tandberg, A. H. S., Ingvarsdottir, A., Aarab, N., Nordtug, T. (2011, October 4-6). *The effects of different oil spill cleanup technologies on body burden and biomarkers in Arctic marine organisms – a laboratory study* [conference proceedings]. Arctic Marine Oil Spill Program Technical Seminar, Banff, Alberta, Canada.

Falfushynska, H., Sokolov, E. P., Haider, F., Opperman, C., Kragl, U., Ruth, W., Stock, M., Glufke, S., Winkel, E. J., Sokolova, I. M. (2019). Effects of a common pharmaceutical, atorvastatin, on energy metabolism and detoxification mechanisms of a marine bivalve *Mytulis edulis. Aquatic Toxicology, 208*, 47-61.

Fernández-Méndez, M., Katlein, C., Rabe, B., Nicolaus, M., Peeken, I., Bakker, K., Flores, H., Boetius, A. (2015). Photosynthetic production in the central Arctic Ocean during the record sea-ice minimum in 2012. *Biogeosciences*, *12*(11), 3525–3549.

Frantzen, M., Falk-Petersen, I.-B., Nahrgang, J., Smith, T. J., Olsen, G. H., Hangstad, T. A., Camus, L. (2012). Toxicity of crude oil and pyrene to the embryos of beach spawning capelin (*Mallotus villosus*). *Aquatic Toxicology, 108,* 42-52.

Gardiner, W. W., Word, J. Q., Word, J. D., Perkins, R. A., McFarlin, K. M., Hester, B. W., Word, L. S., Ray, C. M. (2013). The acute toxicity of chemically and physically dispersed crude oil to key Arctic species under Arctic conditions during the open water season. *Environmental Toxicology and Chemistry*, *32,* 2284-2300.

Geraudie, P., Nahrgang, J., Forget-Leray, J., Minier, C., Camus, L. (2014). *In vivo* effects of environmental concentrations of produced water on the reproductive function of polar cod (*Boreogadus saida*). *Journal of Toxicology and Environmental Health, Part A, 77*(9-11), 557-573.

Gosselin, M., Levasseur, M., and Wheeler, P. A. (1997). Deep Sea Research Part II: Topical Studies in Oceanography – New measurements of phytoplankton and ice algal production in the Arctic Ocean. *Deep Sea Research, 44*(8), 1623-1644.

Grenvald, J. C., Nielsen, T. G., Hjorth, M. (2013). Effects of pyrene exposure and temperature on early development of two co-existing Arctic copepods. *Ecotoxicology, 22*, 184-198.

Hansen B.H., Altin D., Rørvik S. Øverjordet, I., Jager T., and Nordtug T. (2013). Acute exposure of water soluble fractions of marine diesel on Arctic *Calanus glacialis* and boreal *Calanus finmarchicus*: Effects on survival and biomarker response. *Science of the Total Environment, 449,* 276-284.

Hansen, B. H., Altin, D., Olsen, A. J., Nordtug, T. (2012). Acute toxicity of naturally and chemically dispersed oil on the filter-feeding copepod *Calanus finmarchicus*. *Ecotoxicology and Environmental Safety, 86*, 38-46.

Hansen, B. H., Nordtug, N., Altin, D., Booth, A., Hessen, K. M., Olsen, A. J. (2009). Gene expression of GST and CYP330A1 in lipid-rich and lipid poor female *Calanus finmarchicus* (Copepoda: Crustacea) exposed to dispersed crude oil. *Journal of Toxicology and Environmental Health Part A, 72*, 131-139.

Hansen, B. J., Parkerton, T., Nordtug, T., Storseth, T. R., Redman, A. (2019). Modeling the toxicity of dissolved crude oil exposures to characterize the sensitivity of cod (*Gadus morhua*) larvae and role of individual and unresolved hydrocarbons. *Marine Pollution Bulletin, 138*, 286-294.

Hansen, B., Altin, D., Bonaunet, K., Øverjordet, I. (2014). Acute toxicity of eight oil spill response chemicals to temperate, boreal, and Arctic species. *Journal of Toxicology and Environmental Health, Part A, 77*, 495-505.

Hanson, M. and Brain, R. (2021). A method to screen for consistency of effect in laboratory toxicity tests: A case study Anurans and the herbicide atrazine. *Archives of Environmental Toxicology and Chemistry*, *811*, 123-132.

Hanson, M. L. and Brain, R. A. (2020). Context and perspective in ecotoxicology. *Environmental Toxicology and Chemistry, 39*(9), 1655-1656.

Hanson, M. L., Wolff, B. A., Green, J. W., Kivi, M., Panter, G. H., Warned, M. St. J., Ågerstrand, M. Sumpter, J. P. (2017). How we can make ecotoxicology more

valuable to environmental protection. *Science of the Total Environment, 578*, 228-235.

Hanson, M., Baxter, L., Anderson, J., Solomon, K., Brain, R. (2019). Strength of methods assessment for aquatic primary producer toxicity data: A critical review of atrazine studies from the peer-reviewed literature. *Science of the Total Environment, 685*, 1221-1239.

Harris, C. A. and Sumpter, J. P. (2015). Could the quality of published ecotoxicological research be better? *Environmental Science and Technology, 49*(16), 9495-9496.

Harris, C. A., Scott, A. P., Johnson, A. C., Panter, G. H., Sheahan, D., Roberts, M., and Sumpter, J. P. (2014). Principles of sound ecotoxicology. *Environmental Science and Technology, 48,* 3100-3111.

Henry, T. and Pease, A. (2016). ET&C Perspectives: A government perspective. *Environmental Toxicology and Chemistry, 35*(1).

Hjorth, M. and Nielsen, T. G. (2011). Oil exposure in the Arctic: potential impacts on key zooplankton species. *Marine Biology, 158*(6), 1339-1347.

Hobbs, D. A., Warne, M. S. J., and Markich, S. J. (2005). Evaluation of criteria used to assess the quality of aquatic toxicity data. *Integrated Environmental Assessment and Management, 1*, 174-180.

Hsiao, S. C. (1978). Effects of crude oils on the growth of Arctic marine phytoplankton. *Environmental Pollution, 17*, 93-107.

Jensen, M.H., Nielsen, T.G., Dahllöf, I. (2008). Effects of pyrene on grazing and reproduction of *Calanus finmarchicus* and *Calanus glacialis* from Disko Bay, West Greenland. *Aquatic Toxicology, 87*, 99-107.

Kase, R., Korkaric, M., Werner, I., Ågerstrand, M. (2016). Criteria for reporting and evaluating ecotoxicity data (CRED): Comparison and perception of the Klimisch and CRED methods for evaluating reliability and relevance of ecotoxicity studies. *Environmental Sciences Europe, 28*(1), 1-14.

Klimisch, H.L., Andreae, M., Tillmann, U. (1997). A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. *Regulatory Toxicology and Pharmacology, 25*, 1-5.

Küster, A., Bachmann, J., Brandt, U., Ebert I., Hickmann, S., Klein-Goedicke, J., Maack, G., Schmitz, S. (2009). Regulatory demands on data quality for the environmental risk assessment of pharmaceuticals. *Regulatory Toxicology and Pharmacology, 55*, 276-280.

Lazic, S. E., Mellor, J. R., Ashby, M. C., and Munafo, M. R. (2020). A Bayesian predictive approach for dealing with pseudoreplication. *Scientific Reports, 10*, 2366.

Lemcke, S., Holding, J., Moller, E. F., Thyrring, J., Gustavson, K. Juul-Pedersen, T., Sejr, M. K. (2018). Acute oil exposure reduces physiological process rates in Arctic phyto- and zooplankton. *Ecotoxicology, 28*, 26-36.

McGrath, J. A., Fanelli, C. J., Di Toro, D. M., Parkerton, T. F., Redman, A. D., Paumen, M. L., Comber, M., Eadsforth, C. V., den Haan, K. (2018). Re-evaluation of target lipid model-derived HC5 predictions for hydrocarbons. *Environmental Toxicology and Chemistry 37*(6), 1579-1593.

Michel, C., Legendre, L., Ingram, R. G., Gosselin, M., and Levasseur, M. (1996). Carbon budget of sea-ice algae in spring: Evidence of a significant transfer to zooplankton grazers. *Journal of Geophysical Research: Oceans, 101*(C8), 18345–18360.

Moermond, C., Beasley, A., Breton, R., Junghans, M., Laskowski, R., Solomon, K., Zahner, H. (2016). Assessing the reliability of ecotoxicological studies: An overview of current needs and approaches. *Integrated Environmental Assessment and Management, 13*(4), 640-651.

Moore, D., Sibley, P. K., Solomon, K. R. (2016). *Toxicity of salts, metals, and nitrogenous contaminants to cold-water fish under northern conditions*. (Unpublished doctoral dissertation). University of Manitoba, Winnipeg, MB.

Nørregaard, R. D., Nielsen, T. G., Moller, E. F., Strand, J., Espersen, L., Mohl, M. (2014). Evaluating pyrene toxicity on Arctic key copepod species *Calanus hyperboreus. Ecotoxicology, 23*, 163-174.

Office of the Federal Register, National Archives and Records Administration. (2021, October 25). *43 CFR 11 – Natural Resource Damage Assessments.* [Government]. § 11.62 Injury determination phase – injury definition.

Olsen, A. J., Nordtug, T., Altin, D., Lervik, M., Hansen, B. J. (2013a). Effects of dispersed oil on reproduction in the cold water copepod *Calanus finmarchicus* (Gunnerus). *Environmental Toxicology and Chemistry, 32*(9), 2045-2055.

Olsen, G. H., Smit, M. G. D., Carroll, J., Jaeger, I., Smith, T., Camus, L. (2011). Arctic versus temperate comparison of risk assessment metrics for 2-methylnaphthalene. *Marine Environmental Research, 72,* 179-187.

Øverjordet, I. B., Altin, D., Berg, T., Jenssen, B. M., Gabrielsen, G. W., Hansen, B. H. (2014). Acute and sub-lethal response to mercury in Arctic and boreal calanoid copepods. *Aquatic Toxicology, 155*, 160-165.

Pančić, M., Köhler, E., Paulsen, M. L., Toxværd, K., Lacroix, C., Le Floch, S., Hjorth, M., Nielsen, T. G. (2019). Effects of oil spill response technologies on marine microorganisms in the high Arctic. *Marine Environmental Research, 151,* 104785.

Pechenik, J. A. (2006). Larval experience and latent effects – metamorphosis is not a new beginning. *Integrative and Comparative Biology, 46*(3), 323-333.

Petersen, D. G., Reichenberg, F., and Dahllöff, I. (2008). Phototoxicity of pyrene affects benthic algae and bacteria from the Arctic. *Environmental Science and Technology, 42,* 1371-1376.

Philibert, D. A., Philibert, C. P. Lewis, C., Tierney, K. B. (2016). Comparison of diluted bitumen (dilbit) and conventional crude oil toxicity to developing zebrafish. *Environmental Science and Technology, 50*(11), 6091-6098.

Prosser, R. S., Deeth, L., Humeniuk, B., Jeyabalan, T., Hanson, M. (2021). Taxonomic chauvinism in pesticide ecotoxicology. *Environmental Toxicology and Chemistry*, *40*(12), 3223-3225.

R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raimondo, S., Lilavois, C. R., and Barron, M. G. (2015). Web-based interspecies correlation estimation (WebICE) for acute toxicity: user manual. Version 3.3, EPA/600/R-15/192, US Environmental Protection Agency, Office of Research and Development, Gulf Ecology Division. Gulf Breeze, FL.

Rajabi, H., Mosleh, M. H., Mandal, P., Lea-Langton, A., Sedighi, M. (2020). Emissions of volatile organic compounds from crude oil processing – Global emission inventory and environmental release. *Science of the Total Environment, 727*, 138654.

Redman, A. D. and Parkerton, T. F. (2015). Guidance for improving comparability and relevance of oil toxicity tests. *Marine Pollution Bulletin, 98*, 156-170.

Rehberger, K., Kropf, C., Segner, H. (2018). *In vitro* or not *in vitro*: A short journey through a long history. *Environmental Sciences Europe, 30*(23).

Riebell, P. N. and Percy, J. A. (1989, June 7-9). *Acute toxicity of petroleum hydrocarbons to the Arctic littoral mysid,* Mysis oculata *(Fabricus)* [conference proceedings]. Arctic Marine Oil Spill Program Technical Seminar, Calgary, Alberta, Canada.

Rodríguez-Torres, R., Almeda, R., Kristiansen, M., Rist, S., Winding, M. S., Nielsen, T. G. (2020). Ingestion and impact of microplastics on Arctic Calanus copepods. *Aquatic Toxicology, 228*, 105631.

Roth, N., Zilliacus, J., Beronius, A. (2021). Development of the SciRAP approach for evaluating the reliability and relevance of *in vitro* toxicity data. *Frontiers in Ecotoxicology, 3.*

Schneider, K., Schwarz, M., Burkholder, I., Kopp-Schneider, A., Edler, L., Kinsner-Ovaskainen, A., Hartung, T., Hoffman, S. (2009). ToxRTool, a new tool to assess the reliability of toxicological data. *Toxicology Letters, 189*(2), 138-144.

Sonne, C., Dietz, R., Jenssen, B. M., Lam, S. S., Letcher, R. J. (2021). Emerging contaminants and biological effects in Arctic wildlife. *Trends in Ecology & Evolution, 36*(5), 421-429.

Søreide, J. E., Leu, E. V. A., Berge, Jør., Graeve, M., and Falk-Petersen, S. (2010). Timing of blooms, algal food quality and Calanus glacialis reproduction and growth in a changing Arctic. *Global Change Biology, 16*(11), 3154–3163.

Thoré, E. S. J., Steenaerts, L., Philippe, C., Grégoire, A. F., Brendonck, L., Pinceel, T. (2019). Improving the reliability and ecological validity of pharmaceutical risk assessment: Turquoise killifish (*Nothobranchius furzeri*) as a model in

behavioural ecotoxicology. *Environmental Toxicology and Chemistry, 38*(1), 262-270.

Thyrring, J., Juhl, B. K., Holmstrup, M. (2015). Does acute lead (Pb) contamination influence membrane fatty acid composition and freeze tolerance in intertidal blue mussels in Arctic Greenland? *Ecotoxicology, 24*, 2036-2042.

Toxværd, K., Dinh, K. V., Henriksen, O., Hjorth, M., Nielsen, T. G. (2018). Impact of pyrene exposure during overwintering of the Arctic copepod *Calanus glacialis*. *Environmental Science and Technology, 52*, 10328-10336.

Toxværd, K., Dinh, K. V., Henriksen, O., Hjorth, M., Nielsen, T. G. (2019). Delayed effects of pyrene exposure during overwintering on the Arctic copepod *Calanus hyperboreus*. *Aquatic Toxicology, 217*, 105332.

United States Department of the Interior. (2008). *Circum-Arctic Resource Appraisal: Estimates of Undiscovered Oil and Gas North of the Arctic Circle.*

United States Environmental Protection Agency (US EPA). (2003). *A Summary of General Assessment Factors for Evaluating the Quality of Scientific and Technical Information. EPA 100/B-03/001*. https://www.epa.gov/risk/summary-general-assessment-factors-evaluating-quality-scientific-and-technical-information.

United States Environmental Protection Agency (US EPA). (2012). *Ecological Effects Test Guidelines OCSPP 850.4500: Algal Toxicity*. https://nepis.epa.gov/.

United States Environmental Protection Agency (US EPA). (2016). General Ecological Assessment Endpoints (GEAEs) for Ecological Risk Assessment: Second Edition with Generic Ecosystem Services Endpoints Added. EPA/100/F15/05.

United States Environmental Protection Agency. (1994, March). *ECO Update: Using Toxicity Tests in Risk Assessment* (9345.0-05I). Office of Solid Waste and Emergency Response. https://www.epa.gov/sites/default/files/2015-11/documents/v2no1.pdf.

Van der Kraak, G. J., Hosmer, A. J., Hanson, M. L., Kloas, W., Solomon, K. R. (2014). Effects of atrazine in fish, amphibians, and reptiles: An analysis based on quantitative weight of evidence. *Critical Reviews in Toxicology, 44*(S5), 1-66.

Vergauwen, L. (2018). Application of adverse outcome pathways (AOPs) in ecotoxicology – A multi-taxon AOP network to reduce animal testing. *Toxicology Letters, 295*(1), S36-S37.

Wang, B., Yu, G., Huang, J., Hu, H. (2008). Development of species sensitivity distributions and estimation of HC5 of organochlorine pesticides with five statistical approaches. *Ecotoxicology, 17*, 716-724.

Wang, Z., An, C., Lee, K., Owens, E., Chen, Z., Boufadel, M., Taylor, E., Feng, Q. (2021). Factors influencing the fate of oil spilled on shorelines: a review. *Environmental Chemistry Letters, 19*, 1611-1628.

3 A New Toxicity Test with the Arctic Algae *Nitzschia frigida*

3.1 Introduction

The risk to Arctic marine organisms from contaminant exposure is increasing with the growth of human populations, resource development, and shipping in Polar regions globally. Historically in the Arctic, contaminants of concern tended to be transported from southern regions via atmosphere or ocean currents including PCBs, (dichlorodiphenyltrichloroethane (DDT), and other organochlorine compounds (Letcher et al., 2010). With the expanding direct human presence in the region, new types of contaminant concerns are emerging; especially point-source releases. For example, the majority of wastewater treatment in the Arctic takes place via passive methods (e.g., lagoons and wetlands) or by direct discharge into the sea with no treatment (Gunnarsdóttir et al., 2013; Hermanson et al., 2010; Johnson, 2008; Krumhansl et al., 2014), which can contain pharmaceuticals, personal care products, as well as legacy POPs (Chaves-Barquero et al., 2016; Kallenborn et al., 2018; Stroski et al., 2020). Mineral extraction and hydrocarbon exploration also pose a threat through both the deliberate discharge of mining effluents into the sea and the potential for accidental oil spills when exploiting these resources (AMAP, 2007; Tolvanen et al., 2019). Of concern is the limited availability of high-quality laboratory toxicity data in the literature using Arctic test species to assess the risks posed by these contaminants. In Chapter 2 of this thesis, we found that only 18% of laboratory toxicity tests using Arctic species could be considered reliable for risk assessment purposes, and recommended expanded efforts to characterize toxicity to Arctic species specifically.

Data generated through standard toxicity tests are used to elucidate environmental risk of contaminant exposure under the ERA framework (Health Canada, 2018). These test methods are standardized and validated, which allows data to be comparable between laboratories and over time (ECCC, 2013). They typically use temperate species as test organisms such as the marine algae *Skeletonema costatum* (diatom) due to their ease of culturing and representativeness of many habitats; however, critical differences in the structure and function of Arctic and temperate ecosystems, as well as physiological differences between the organisms themselves,

may prevent these data from being readily extrapolated to Arctic risk assessment without significant uncertainty (Chapman and Riddle, 2005; US EPA 1994; US EPA 1996). SSDs and hazard concentrations for Arctic and non-Arctic species have been generated; however, results have not uniformly indicated that one group is more or less sensitive than the other (e.g., Bejarano et al., 2017; Camus et al., 2015; Hansen et al., 2014; Olsen et al., 2011). Further, the potential for recovery after contaminant exposure has not been thoroughly addressed, leaving a marked gap in scientific knowledge regarding long-term, ecosystem-level effects. This lack of clear, repeatable evidence comparing Arctic and non-Arctic species' sensitivities highlights the need for reliable data for use in ERA within the polar North.

Ice-associated organisms (e.g., ice algae) could be at greater risk from local point source releases of contaminants such as oil from spills or leaks, as the buoyancy of these contaminants combined with a lack of wind-driving mixing of the water column may result in their being trapped in the under-ice environment (Chapman and Riddle, 2005; Cota et al., 1991). The low temperatures and constant dark or light polar seasons may influence the behaviour and characteristics of contaminants, and differences in lipid content and metabolic rates of Arctic species may result in unexpected effects from exposure (Chapman and Riddle, 2005). Additionally, relatively low species diversity in the Arctic contributes to a lack of functional redundancy within the ecosystem, resulting less potential for recovery after contaminant exposure (Aune et al., 2018).

Algae are an important component in the Canadian Arctic food web and play key functional roles in the ecosystem. During the spring melt when food is scarce, the annual ice algae bloom is a food source critical to the survival of many species as it thrives attached to the ice, and later to zooplankton and other pelagic and benthic grazers as cells begin descending through the water column and settle on the ocean floor (Bradstreet and Cross, 1982; Darnis et al., 2012; Kohlbach et al., 2017; Michel et al., 1996; Mundy and Meiners, 2021; Nuttall et al., 2005; Søreide et al., 2010). These organisms support the biological success of the other members in the Arctic food web, for which this fat-rich pulse of fuel and carbon drives almost all other functions within this ecosystem (Artamonova et al., 2017a, 2017b; Gosselin et al., 1997; Kohlbach et al.,

2016; Michel et al., 1996; Søreide et al., 2010). These species are ecologically relevant and warrant investigation as possible test species for toxicity studies.

Aside from their ecological relevance, algae are an excellent group of organisms to prioritize in toxicity testing due to their sensitivity to many substances and their ability to be cultured year-round in the laboratory (Staveley and Smrchek, 2005). Widely accepted standard methods for algal tests have been put forth by ECCC (ECCC, 2007), OECD (OECD, 2011), US EPA (US EPA, 2012), NIWA (NIWA, 1998), and ASTM (ASTM, 2012). The general principles of these methods are fundamentally similar; algae are cultured in nutrient rich media, typically under continuous light to obtain maximum growth within the test period, and exposed to contaminants for a sufficient duration to observe growth inhibition in test vessels when compared to controls. Toxicity is assessed based on a variety of endpoints (i.e., specific growth rate, fold increase, cell yield) to provide data for risk assessors to use in determining whether environmental concentrations of contaminants present significant risk to organisms.

The overall objective of this paper was to further the understanding of Arctic species' sensitivity to contaminants in relation to their temperate counterparts to best inform ERA as it relates to primary producers. This was accomplished by 1) optimizing culture conditions for the laboratory toxicity testing of candidate Arctic algal species; 2) developing a protocol for toxicity testing with the selected candidate species, and 3) contrasting the new species with the standard marine test species *S. costatum*. The resulting new protocol and the standard protocol were performed three times to assess its repeatability with three different test substances (copper, zinc, and the IC 1-methylnaphthalene) and test the hypothesis that Arctic ice algae is more sensitive to these compounds than the standard test species *S. costatum*.

3.2 Methods

3.2.1 General Considerations

3.2.1.1 Test species selection and sources

A total of three candidate species were evaluated as potential Arctic test organisms: *Nitzschia frigida*, *Melosira arctica*, and *Porosira glacialis*. They were

selected due to their availability in culture collections, as well as predominance and contribution to function in Arctic ecosystems. All are ice-associated species that dominate the ice-water interface in the Arctic Ocean (Artamonova et al., 2017a; Aumack and Juhl, 2015; Poulin et al., 2014).

N. frigida is a pennate diatom that grows in branching colonies attached to sea ice (Figure S18). It is widely prevalent in communities within the bottom layer of first-year sea-ice during the spring, comprising up to 85% of all cells sampled in some studies (Michel et al., 1996, 2002; Riedel et al., 2003; Rózańska et al., 2009). *M. arctica* is a centric diatom that can often be observed in strands up to 30 cm long attached to the bottom of sea-ice (Figure S19). It is frequently found to be one of the dominant species in Arctic sea ice in the late spring and early summer, often succeeding communities of *N. frigida* (Gosselin et al., 1997; Poulin et al., 2014; Riedel et al., 2003). Additionally, this species may contribute up to 90% of all NPP in ice algal communities where they are dominant (Fernández-Méndez et al., 2014; Poulin et al., 2014). *P. glacialis* is another common Arctic diatom of importance in the Arctic Ocean (Artamonova et al., 2017a, 2017b). This species was selected for its role in the Arctic ecosystem, and because it had been previously used in toxicity testing (e.g., Camus et al., 2015) and had some amount of life history background available in the literature (e.g., Artamonova et al., 2017a, 2017b; Ha et al., 2014; McMinn et al., 2005). Neither *N. frigida* nor *M. arctica* had been previously used as toxicity test organisms in the publicly available literature at the time of this writing, although they have been minimally studied in other contexts.

Cultures of *N. frigida* and *M. arctica* were unavailable to purchase from commercial culture collections but were available from cultures originally isolated by Dr. Andrew Juhl in May 2011 from melted sea ice samples collected near Barrow, Alaska (Aumack and Juhl, 2015). Cultures of both species were monoalgal and non-axenic. *P. glacialis* was commercially obtained as a monoalgal, axenic culture from the Bigelow National Center for Marine Algae and Microbiota in East Boothbay, Maine (CCMP668). *S. costatum* is a centric, chain-forming diatom that is a standard species for algal toxicity testing (i.e., ISO, 2016; US EPA, 2012). It has a cosmopolitan distribution but thrives in temperate environments at optimal temperatures of 25°C (Kooistra et al.,

2008; University of British Columbia, 2012). As a standard species it is representative of a large proportion of toxicity data against which Arctic species data can be compared.

3.2.1.2 Glassware and aseptic technique

All laboratory glassware used for algal culturing and toxicity testing was cleaned prior to use according to procedures that adhere to guidelines found in standard methods (i.e., those of the US EPA, OECD, ISO, ECCC). Briefly, this includes washing in hot tap water with mild lab detergent (e.g., Liquinox) followed by a thorough rinse in 15% hydrochloric acid (Fisher Scientific, Chemical Abstracts Service [CAS] Registry Number 7647-01-0), three rinses with hot tap water, and three rinses with distilled water (prepared in-house using a Thermo-Scientific Barnstead Mega-Pure Still MP-1 distiller). Glassware was then dried in a drying oven (Fisher Scientific Isotemp Oven) at 60ºC for at least one hour. All media was autoclaved in pre-washed flasks outfitted with cotton-filled cheesecloth bungs prior to inoculation with algae. All other laboratory equipment (e.g., pipette tips, microcentrifuge tubes, test tubes) were sterile before use in algal applications, either by the manufacturer or by autoclaving for 15 minutes at 121ºC (Tuttnauer Benchtop Autoclave 3850E-B/L).

All procedures that could compromise the sterility of algal cultures (e.g., transferring stock cultures, inoculating EUs, preparing algal suspensions) were performed using aseptic technique (Food and Agriculture Organization of the United Nations, 2004). This includes working in a biological safety cabinet (REDISHIP Purifier Logic+ A2 Biosafety Cabinet, 302389001) sprayed with 70% ethanol (Fisher Scientific, CAS 64-17-5) and UV-sterilized for at least 10 minutes prior to use, flaming the necks of vessels with a propane torch prior to opening in order to create air convection currents away from the contents, and ensuring all tools are properly sterilized before use.

All test solutions and culturing media were prepared using seawater (or artificial saltwater, see Section 2.1.4 Culture media) that had been filter-sterilized using a bottle-top filtration unit (Celltreat bottle top filter, 1000 mL, 0.22 µm PES filter) prior to use.

3.2.1.3 Initial culturing

Upon receipt, flasks of *N. frigida* and *M. arctica* were immediately placed in culture conditions recommended by Dr. Juhl for 48 hours before any transfers occurred. Briefly, these conditions were 3 ± 1ºC under a continuous photoperiod of full-spectrum fluorescent light at ~60 µmol photons m^{-2} s^{-1} in a refrigerator (Hamilton Beach 19" 2.7 ft^3 compact refrigerator, #REFHB270BE) outfitted with a high-output full-spectrum fluorescent light (SunBlaster T5 8W 3500K) and digital temperature controller (Inkbird digital temperature controller, ITC-308-11). They were then transferred into autoclaved 125-mL Erlenmeyer flasks with cotton-filled cheesecloth bungs containing 50 mL of fresh, sterile Guillard's f/2 media with silicate (Table S6, Guillard and Ryther, 1962; Guillard, 1975), prepared with artificial seawater (ASW; Instant Ocean Sea Salts, SS15-10, and Milli-Q distilled water). Cultures were incubated under these conditions for two weeks and were again transferred into new 125-mL flasks containing fresh media. From this point forward, these parent cultures were serially transferred bi-weekly into new 125-mL or 250-mL flasks (depending on volume requirements) to maintain a stock culture of cells in exponential growth. All stock transfers were performed according to the HMSC laboratory method for this task.

The available culture of *N. frigida* that was used in this study was monoalgal but non-axenic. Axenic cultures are typically required in toxicity testing so that any effects observed can be more definitively attributed to the compound of interest and confounding due to bacterial contamination can be ruled out. However, it is also well-known that many species of phytoplankton have symbiotic relationships with bacteria, whether they be mutualistic, parasitic, or commensal (Croft et al., 2005; Meng et al., 2020) and *Nitzschia* species in particular have demonstrated their reliance on these relationships in certain contexts (Sison-Mangus et al., 2014, Stewart et al., 1997). Due to the limited available knowledge regarding the relationships between bacteria and *Nitschia frigida* specifically, it was decided to purify the cultures only to the point that cells were not so aggregated that it became impossible to count with a hemocytometer. Purification was first attempted using the single-cell isolation by micropipette technique outlined in Guillard 2005. Following a 30-day incubation period, visible aggregates of

bacteria continued to be observed, but with significantly less cell aggregation. An attempt to further purify the culture via antibiotics was used following a protocol modified from Kan and Pan (2010); however, the cultures had not yielded sufficient growth after four weeks of incubation and were discarded.

M. arctica cultures began to fail shortly after receipt with reproduction of cells ceasing regardless of modifications to light intensity. Parent cultures were deemed non-viable on 5 July 2019. Due to the initial difficulties with culturing this species and the long durations required to observe differences in growth rates the decision was made to halt work on this candidate species. *N. frigida* and *P. glacialis* were carried forward to determine the optimal culture media, counting method, and test vessel requirements for use in toxicity testing.

3.2.1.4 Culture media

In order to determine an optimal media type that supported the growth of these species and was both practical and cost effective, a total of three media types were tested. These included Harrison's, Guillard's f/2 prepared from scratch, and Guillard's f/2 prepared from pre-formulated packages, each prepared with both natural and ASW, for a total of six combinations. The chemical composition of each is presented in Table S6, with the exception of the pre-packaged Guillard's f/2 media. This is a pre-formulated dry mixture of f/2 available for purchase from commercial suppliers (Micro Algae Grow Mass Pack, Florida Aqua Farms, FA-MP1-100). It is often used to prepare large volumes of media to sustain algal cultures used as an industrial food source for shellfish. Distilled water was added to the pre-packaged formulation to create a concentrated nutrient solution containing all required compounds except silicate. This solution was autoclaved to preserve shelf life for storage and was then added to 0.22-µm filtered NSW or ASW at a ratio of 1 mL : 1 L. Silicate was prepared separately, autoclaved, and added aseptically at the same ratio of 1 mL : L after the media was autoclaved to avoid precipitation. This product is cost-effective to purchase and significantly reduces the time and labour required to prepare media from scratch. However, the mixture is a blend of technical and reagent grade ingredients, which means reduced purity and clarity of the media when used in benchtop toxicity testing.

The f/2 media prepared from scratch and Harrison's media were made by creating multiple concentrated nutrient solutions and adding varying amounts to 0.22-µm natural or artificial filtered seawater. Nutrient solutions were prepared by adding reagent grade chemicals to distilled water and autoclaving. ASW was prepared by adding Instant Ocean Sea Salts to distilled water to achieve a salinity of 30-35 psu. Solution was prepared > 24 hours prior to use to ensure all salts had completely dissolved.

3.2.1.5 Counting methods

For all work prior to the test vessel optimization trial (i.e., all tests conducted with *N. frigida* and *P. glacialis*) cell counts were accomplished by using a hemocytometer. Briefly, this included counting all cells in all nine 1 x 1 mm squares of a hemocytometer (Hausser Scientific Bright-Line Hemacytometer, Z359629) and transforming the raw count to achieve density in cells/mL. For each sample, two hemocytometers (four chambers) were loaded and counted and the average value was used for analysis. Due to the colonial nature of *N. frigida,* multiple aggregates of cells often overlap in the field of view rendering it difficult to count. To address this, several attempts were made to break apart large aggregates through sonication (Intertek Ultrasonic Cleaner, CD-4860; 60 Hz; ten, thirty, and sixty minutes) though they ultimately proved unsuccessful and aggregates remained intact. It was then discovered that cultures that had been swirled manually once daily resulted in visibly less aggregation than those that were left untouched for longer periods of time, so this practice was implemented moving forward to facilitate enumeration; however, cultures used to prepare standard curves were still too aggregated to easily assess microscopically. In these cases, the aggregated samples were thoroughly mixed and the volume doubled with sterile media to reduce overlap of cell aggregates in the field of view. This process was repeated until cells could be accurately quantified, and a dilution factor was applied to raw counts to achieve initial cell density in cells/mL. This process was time-consuming and inefficient, and variability at low concentrations was greater for *N. frigida* than for *S. costatum* as presented in Section 3.3 Results. Therefore, two alternative counting methods for estimating cell density were then assessed for their efficiency and accuracy.

Serial dilutions of algal cultures for each species used in definitive toxicity testing (*N. frigida* and *S. costatum)* were prepared and concentrations confirmed with manual counts on a hemocytometer as described above. Samples of each dilution were run on a Coulter counter (Beckman Coulter Multisizer 4e, 100 µm aperture, control mode: volumetric, 100 µL) and a fluorometer (Horiba Scientific Aqualog Fluorometer, 2D emission, excitation wavelength 340 nm, emission wavelength ~681 nm). Excitation and emission wavelengths were initially selected based on chlorophyll fluorescence values from the literature (e.g., Simis et al., 2012) and refined based on the intensity of fluorescence within this range when run using 3D emission. Conditions were the same for both tested species. All data were normalized for day-to-day comparisons by dividing the relative fluorescence units obtained in analysis by the area of the Raman scattering peak produced that day. Standard curves were created, and the process was repeated three times for each species. The r^2 values were calculated with 0.9 being assigned as the minimum acceptable value to deem the instrument a valid surrogate for manual counts.

The standard curves created using the Coulter counter and fluorometer met this minimum criterion for both species used in toxicity testing (Figure S20-31). However, other drawbacks for each instrument became apparent throughout this process. These drawbacks may not be an issue for studies where cell counts are completed only at the end of exposure, but these experiments required sampling of live EUs at multiple time points with vessels being incubated again after sampling, which introduced further limitations from these instruments. Firstly, the Coulter counter lacks the ability to differentiate between live and dead cells as it counts all particles that pass through the aperture. For algicidal test substances these data would be insufficient for analysis. Also, the minimum volume requirement for samples to be analyzed with the Coulter counter was 20 mL, which exceeds the EU volume used in many standard tests (e.g., ECCC, 2007; NIWA,1998). If growth was sufficient in test vessels, small volumes could be withdrawn from active units and diluted prior to running to achieve sufficient volume; however, for units with limited growth (e.g., units with high concentrations of a growth-inhibiting compound) this would likely not be attainable. Further, the aperture of the instrument must be inserted into the sample to take the measurement. Since it is difficult

to sterilize this piece of lab equipment to avoid contaminating an ongoing toxicity test, samples of EUs would again need to be removed from the test to be analyzed and subsequently discarded. This means that test volumes would have to be greatly increased, likely to the point where space and resource use would not allow successful use of the Coulter counter as an analytical tool for this purpose. The lowest algal concentrations that was able to be reliably quantified on the Coulter counter based on variability comparisons with manual hemocytometer counts (i.e., Coulter counter values within 20% of manual counts) was 8,862 cells/mL, establishing this value as the limit of detection (LOD) for the instrument. Since the variability of algal density in inoculum controls (ICs) extends below this range (Figure S32), these units would still require manual counting with a hemocytometer to be accurate.

The fluorometer proved to be a good surrogate for manual cell counts for both species initially based on a linear relationship and acceptable r^2 value of standard curves (Figures S20-27). However, the fluorometer demonstrated less reliability in definitive trials for *S. costatum* only (i.e., no linear relationship between manual cell counts and fluorometry values), while the *N. frigida* relationship remained reliable (i.e., linearity between manual counts and fluorometry values). Reasons for this discrepancy were explored with the only difference between samples being the algal species and background matrix (commercial formulation of Guillard's F/2 for *S. costatum* and HESNW for *N. frigida*). Though the same media type was used to create initial curves with *S. costatum* as was used in the definitive trials, the use of a new batch of media may have contributed to this issue. Troubleshooting steps, including restarting the instrument, cleaning the cuvettes and re-loading samples, did not address the issue. As the trial was already underway, troubleshooting other batches of growth media to use in exposure was no longer an option without restarting the test; therefore, manual cell counts were selected as the method of assessing growth for this species. For consistency between trials, the fluorometer was used as a surrogate for cell counts for *N. frigida* but not for *S. costatum* for all definitive tests. Additionally, when the three standard curves created for *N. frigida* were combined into a single figure, the relationship became very poor (Figure S23). This day-to-day variation was accounted for in *N. frigida* test assessment by creating standard curves for each day samples were

run. From the range-finding trial forward, serial dilutions of at least five concentrations were prepared using the parent culture from the trial to be analyzed. All five concentrations were assessed by fluorometry at the beginning of the series of samples, and the high, medium, and low concentrations were re-run every fifteen samples thereafter. These serial dilutions were counted on a hemocytometer as described above and used to create standard curves where the formula was then applied to the fluorometry values to obtain values in cells/mL. The lowest algal concentration the fluorometer was able to detect with accuracy when compared to manual cell counts (i.e., within 20% of manual count) was 18,944 cells/mL; therefore, this value was established as the lower LOD for this instrument and all ICs for all species and trials were counted manually with a hemocytometer. All *S. costatum* units were counted manually with a hemocytometer and required no modifications to the procedure as cells were easily distinguishable.

3.2.2 Optimization of Growth Conditions

3.2.2.1 General test methods

Test cultures and experiments were maintained in a walk-in environmentally controlled chamber set to the appropriate temperature depending on the species. Cultures were confirmed to be in exponential growth at test initiation by plotting a growth curve over three to seven (temperate species) or fourteen (Arctic species) days prior to launching each experiment. Cultures were sampled every 24 hours (± 1 hour) by aseptically removing 1 mL of well-mixed culture and counting with a hemocytometer. All trials included four replicates of each treatment (three EUs, plus a surrogate of each treatment for measuring pH at the beginning and end of the test).

For each species five ICs were also included to confirm the number of cells inoculated at the start of the experiment, which can help account for day-to-day variation in pipetting or technique. These were surrogate units placed randomly among the EUs during the algae allocation component of the test and were preserved with 1% glutaraldehyde (Sigma-Aldrich, CAS 111-30-8) immediately after inoculation. They were sampled and counted with a hemocytometer to determine the actual average initial cell density for each trial from which several growth endpoints were calculated. Differences

in precision and accuracy of this technique for both species is discussed in Results Section 3.2.2.2 and demonstrated in Figure S32. Treatments were randomly assigned EU numbers by using the =RAND() function on Microsoft Excel and re-randomized in their test tube racks once (*S. costatum)* or twice (*N. frigida*) daily to account for room and location effects.

Cell densities in experimental cultures were adjusted by the addition of a calculated volume of sterile media to achieve the appropriate density for inoculation (10,000 cells/mL for *N. frigida* and *S. costatum;* 5000 cells/mL for *P. glacialis*). Due to the relatively large size of *P. glacialis* cells (20-40 µm), initial cell density was halved for this species as was done in Camus et al., 2015. All inoculum volumes were ≤ 10% of the final volume in each EU such as not to dilute the test solution.

Water quality (temperature, dissolved oxygen [DO] saturation, salinity, and pH) were measured in test solutions at the start of each exposure, and pH was measured in pH surrogate units at the start and end of each exposure. At these time points, temperature and DO saturation were measured with a YSI ProODO optical dissolved oxygen instrument and salinity and pH were measured with a YSI MultiLab 4010-2 digital meter. Temperature was continuously monitored throughout all experiments with a logging thermometer (Fisherbrand™ Traceable™ thermometer, 15-077-976) and instantaneous and 24-hour historical minimum/maximum temperatures were manually recorded daily. Photosynthetically active radiation (PAR, wavelengths of 400-700 nm) was measured daily using a handheld full-spectrum quantum meter (Hoskin Scientific, MQ-501) and placing the sensor among the test vessels at the same level as the liquid in the flasks.

EUs were manually swirled once (Arctic species) or twice (temperate species) daily until all cells were in a homogenous suspension to avoid cell adhesion to vessel walls, and units were re-randomized in the exposure area at this time to account for variations in temperature or light penetration.

At the end of each test period, cell division was arrested by the addition of 1 mL of 1% glutaraldehyde. EUs were sealed with Dura-Seal to prevent evaporation during the assessment process. Cell densities in each EU were determined by counting with a hemocytometer as described previously unless otherwise stated. Growth endpoints

were calculated at the end of each exposure period for all tests and can be found in Table 3.1.

Table 3.1 Growth endpoints calculated from 8-, 10-, 12-, 14-, and 16-day exposures of *Nitzschia frigida* to copper, zinc and 1-methylnaphthalene.

Growth Endpoint	Equation
Yield	$b_j - b_i$
Fold increase	$b_j \div b_i$
Growth rate	$K' = \frac{\ln(b_j - b_i)}{t}$
Average specific growth rate	$\bar{r}_{i-j} = \frac{\ln(b_j) - \ln(b_i)}{t}$
Divisions per day	$div.day^{-1} = \frac{K'}{\ln(2)}$

K'= growth rate of observed cell density from time *i* to time *j*.
$\bar{r}_{i-j}$ = average specific growth rate per day of observed cell density from time *i* to time *j*.
b_i = observed cell density at beginning of the observation interval (time *i*)
b_j = observed cell density at the end of the observation interval (time *j*)
t = time interval from *i* to *j* in days

3.2.2.2 Media optimization trials

A total of six treatments were used in the media optimization trial (Harrison's with ASW [HESAW], Harrison's with NSW [HESNW], f/2 prepared from reagent-grade chemicals with ASW, f/2 prepared from reagent-grade chemicals with NSW, f/2 prepared from pre-formulated packages with ASW, f/2 prepared from pre-formulated packages with NSW). Flasks were incubated in a controlled environmental chamber under conditions recommended by the culture curator for the estimated duration to reach exponential growth.

3.2.2.3 Test vessel optimization trials

A total of three test vessel types were examined in the test vessel optimization trial: 125-mL Erlenmeyer flasks outfitted with cotton-filled cheesecloth bungs (e.g., US EPA, 1971); 25-mL borosilicate glass test tubes with clear caps (e.g., OECD, 2011);

and a polystyrene 96-well cell culture plate (e.g., ECCC, 2007). There were no ICs for the microplate units as cell density can be measured in active EUs using the microplate method. This was completed each day of the trial. The trial was conducted under the same environmental conditions as the media optimization trial.

Cell densities for the microplate units were assessed every 24 h (± 1 h) using a microplate spectrophotometer (BioTek Epoch 2 Microplate Spectrophotometer; excitation wavelength 340 nm). Standard curves were prepared as previously described for determination of cell densities in test tubes and flasks.

3.2.2.4 Temperature and light trials

To determine the maximum temperature that would be tested, a preliminary trial was conducted where 50 mL each of *N. frigida* and *P. glacialis* were incubated under continuous light at 90 µmol photons m^{-2} s^{-1} PAR. Initial temperature was set to 2ºC and temperature was increased by 2ºC per day until 16ºC was reached. Flasks were sampled daily and counted with a hemocytometer. Though variable, results indicated that cell growth stopped increasing at 16ºC for both species (Figures S33-34). Therefore, 14ºC was selected as the highest temperature that would be tested. Trials were then conducted in a full-factorial design at temperatures of 2ºC, 6ºC, 10ºC, and 14ºC and light intensities lower than (60 µmol photons m^{-2} s^{-1}), equal to (90 µmol photons m^{-2} s^{-1}), and higher than (120 µmol photons m^{-2} s^{-1}) values typically found in a standard algal toxicity test. Trials of the same temperature were conducted simultaneously in a controlled environmental chamber. Light intensity over the exposure area was set to the highest value to be tested and cheesecloth was used to shade two sections to achieve the middle and low light intensities.

The limited growth of *P. glacialis* in any of the tests conducted up until this point resulted in discounting it as a candidate species and *N. frigida* was the only Arctic species tested moving forward.

3.2.3 Toxicity Testing

3.2.3.1 General test methods

The final test condition recommendations from the optimization trials above can be found in Section 3.1. These were then implemented for all range-finding and definitive tests to compare the relative sensitivity between a standard temperate species and an Arctic species.

EU volume was doubled to 20 mL for the range-finding and definitive tests to accommodate the aseptic removal of small volumes from each unit on four to five occasions throughout the test. This occurred so that EC50 values could be calculated from at least four different time points to calculate incipient EC50 values (i.e., $EC50_{min}$). This sampling was not conducted for the range-finding test, but the volume was doubled to ensure that similar growth was achieved in twice the volume before proceeding with the change for the definitive tests. An additional temperature monitoring device was used throughout the range-finding and definitive trials. A HOBO TidbiT v2 Temperature Data Logger (Hoskin Scientific, UTBI-001) was deployed at test initiation and placed in a 50-mL beaker of seawater in the exposure area to measure and log temperature every sixty seconds. These temperature profiles can be found in Figure S35.

Three replicates of each treatment plus a surrogate unit of each treatment, included to measure pH at the beginning and end of the exposure period without contaminating active EUs, were included in all tests. Surrogate pH units were aseptically sampled and counted with a hemocytometer daily (*S. costatum*) or three times per week (*N. frigida*) track trial progress. This ensured that the test could be terminated quickly if sufficient growth was not achieved to conserve time and resources.

N. frigida tests included 16 days of exposure followed by 14 days of recovery, with definitive data for EC50 value comparisons with *S. costatum* being collected at 14 days post-exposure. *S. costatum* tests included 7 days of exposure followed by 4 days of recovery, with definitive data for EC50 value comparisons with *N. frigida* being collected at 4 days post-exposure (96 hours).

Cell density in each EU was determined by counting on a hemocytometer (*S. costatum* units) or by fluorometry (*N. frigida*) as described in Methods Section 2.1.5. Growth endpoints were calculated for each species and test as previously described.

3.2.3.2 Range-finding tests

A total of three concentrations of each reference toxicant (copper sulphate, zinc sulphate, and 1-methylnaphthalene) were tested using both *N. frigida* and *S. costatum.* Concentrations for both species were selected based on a range that captures many EC50 values from the literature, and on in-house data previously collected using *S. costatum.* The nominal concentrations were 1000, 2000, and 4000 µg/L Cu or Zn and 310, 2820, and 16,920 µg/L 1-methylnaphthalene.

3.2.3.3 Definitive tests

The first definitive test with *S. costatum* failed due to insufficient growth in the control units, and bacterial contamination was observed in the parent cultures. Due to the time-consuming process of purifying algal cultures, a different culture of the same species was selected to continue the testing process. This culture was obtained from the Department of Fisheries and Oceans Centre for Offshore Oil, Gas, and Energy Research Collection (DFO COOGER).

Definitive tests were conducted similarly to range-finding tests with a few exceptions. Briefly, five concentrations of each compound were tested on each species. Nominal concentrations of copper for *S. costatum* tests were 1320, 1650, 2063, 2578, and 3223 µg/L in all but one test where a calculation error resulted in concentrations of 1650, 2063, 2578, 3223, and 4028 µg/L; zinc concentrations were 1650, 2063, 2578, 3223, and 4028 µg/L; and 1-methylnaphthalene concentrations were 100, 310, 940, 2820, and 8460 µg/L. Nominal concentrations of both copper and zinc for *N. frigida* tests were 450, 675, 1013, 1519, and 2278 µg/L, and 1-methylnaphthalene concentrations were 30, 100, 310, 940, and 2820 µg/L. Additionally, O-ring controls were introduced in the second trial to account for the physical presence of the silicone O-rings in the methylnaphthalene units. These test tubes with O-rings were identical to

methylnaphthalene EUs except the O-rings were not treated with the compound. Test duration was also extended from 4 to 7 days for *S. costatum* and from 14 to 16 days for *N. frigida* to ensure sufficient differences in growth could be achieved for calculating the $EC50_{min}$; however, samples were still taken at the initially targeted time points so that ECx values could be calculated (96 hours for *S. costatum* and 14 days for *N. frigida*). Samples from every EU were also taken at various other time points throughout the exposure (days 5, 6, and 7 for *S. costatum*; days 8, 10, 12, and 16 for *N. frigida*) for calculating $EC50_{min}$ values. Lastly, after 96 hours (*S. costatum)* or 14 days (*N. frigida*) of exposure, subsamples of each unit were washed and re-suspended in clean media to observe the potential for recovery after growth inhibition as described below.

3.2.3.4 Recovery tests

The recovery method used for these trials was adapted from Baxter et al., 2014. Briefly, 4 mL of well-mixed algal suspension was aseptically removed from each EU at the end of exposure and transferred to 2 x 2-mL sterile microcentrifuge tubes. Samples were pelletized using a benchtop centrifuge (Eppendorf microcentrifuge 5424-R) equipped with an 11-cm rotor (Eppendorf, FA-45-24-11) at 5000 *g* for 20 minutes. 1.75 mL of supernatant was removed from each tube and 1.75 mL of fresh, sterile media was added. Tubes were centrifuged again, supernatant removed, and fresh media added. Tubes were then inverted ten times or until the pellet became re-suspended in the liquid media, then all tubes of the same treatment were pooled into a 50-mL glass beaker. Each beaker of pooled replicates was sampled and counted with a hemocytometer to determine cell density, which was then used to calculate the volume required to inoculate new cultures in clean media for recovery assessment. Recovery tests were conducted identically to range-finding tests except no test substance was used and surrogate pH units were not counted during the exposure.

3.2.3.5 Test substances and exposure confirmation

Reagent grade copper sulphate pentahydrate (Sigma-Aldrich, ≥ 98.0%, CAS 7758-99-8) was dissolved into distilled water to achieve a stock solution with a nominal

concentration of 2.000 g/L $CuSO_4$ (measured concentration of 586 mg/L copper; 29% of the nominal compound concentration). Reagent grade zinc sulphate (Honeywell Fluka, ≥99.5%, CAS 7446-20) was dissolved into distilled water to achieve a stock solution with a nominal concentration of 2.577 g/L $ZnSO_4$ (measured concentration of 612 mg/L zinc; 24% of the nominal compound concentration). Stock solutions were then spiked into their respective media to create test solutions for each trial.

Copper and zinc stock solutions that were diluted for testing were analyzed by the Research and Productivity Council (RPC) in Fredericton, New Brunswick on 25 August 2020 (reporting limit = 0.01 mg/L) and 12 November 2020 (reporting limit = 0.001 mg/L), respectively, in accordance with the US EPA 200.7/200.8 methods. These measured concentrations were used to calculate volumes required for the preparation of test solutions. Individual test solutions from all definitive trials were sent to RPC for analysis on May 4, 2021. These included test solution samples collected at trial initiation and pooled replicates from trial termination (Figures S36-38). However, due to an error in sample storage conditions, nominal concentrations calculated from the measured stocks were used to calculate EC50 values for both metals (see Discussion for further information). Concentrations of metals in controls for all trials were lower than the sum of expected background concentrations in seawater plus the known concentration in the media.

1-methylnaphthalene concentrations in EUs were achieved via a passive dosing system using polydimethylsiloxane (PDMS) O-rings (Silicone O-Rings West; diameter 2.5 cm, thickness 2.33 mm, mass ~1.06 g; AS568B-210) that was established by Butler et al. (2016) and Renegar et al. (2017). Briefly, O-rings were cleaned by soaking in ethyl acetate (Sigma-Aldrich, 99.8%, CAS 141-78-6) for 24 hours followed by at least 24 hours of soaking in high-performance liquid chromatography (HPLC) grade methanol (Fisher Scientific, CAS 67-56-1) which was refreshed three times throughout this period. Then, O-rings were added to a stock solution of 1-methylnaphthalene (Acros Organics, 96%, CAS 90-12-0) and HPLC grade methanol in 125-mL Erlenmeyer flasks and placed on an MaxQ™ 2000 benchtop orbital shaker (Thermo Scientific, SHKE2000) at 200 revolutions per minute (RPM) for 72 hours. O-rings were then rinsed three times in distilled water, patted dry with Kimwipes, and stored in aluminum foil packages until

use. Two O-rings of the appropriate concentration were placed at the bottom of each 25-mL test tube containing 20 mL media and equilibrated on an orbital shaker at 125 RPM for at least 24 hours prior to the start of exposure. Samples were taken from surrogate test vessels at test initiation and test termination for chemical analysis.

1-methylnaphthalene concentrations were confirmed in-house using a Horiba Aqualog fluorometer. Samples were run using 2D emission, excitation wavelength of 265 nm and emission wavelength of ~332 nm. Standard curves were generated by plotting fluorometry values against measured concentrations that were independently validated by a third-party laboratory (Bureau Veritas, Bedford, Nova Scotia) using gas chromatography-mass spectrometry (GC-MS) and the US EPA 8270 analytical test method to obtain readings in µg/L. All data were normalized for day-to-day comparisons by dividing the relative fluorescence units obtained in analysis by the area of the Raman scattering peak produced that day. Concentrations of 1-methylnaphthalene in controls were deemed to be zero as the values produced by fluorometry were below the LOD for the instrument, which was deemed to be 22.5 µg/L of 1-methylnaphthalene by way of comparing background noise to measured values.

3.2.4 Statistical Analyses

Statistical analyses were performed and figures created in R (R Core Team, 2021). Negative values (e.g., fluorometry values below the LOD, indicating no growth) were replaced with zeros for statistical analysis purposes. Additionally, a single anomalous cell yield value was removed from analysis as there was little likelihood that it was possible to obtain and was more likely due to machine or operator error (cell count of ~6 million cells/mL).

Assumptions of normality were assessed using Shapiro-Wilk's test and homogeneity of variance using Bartlett's test. Statistically significant differences ($p < 0.05$) were determined using one-way analysis of variance (ANOVA) and Dunnett's post-hoc test to determine where treatments were different from controls when data were normally distributed. Non-normally distributed data with unequal variances were assessed for statistical differences by way of the Kruskal-Wallis rank sum test. Statistical differences between groups that were not compared to controls (e.g., trials

conducted during the exposure conditions optimization phase) were calculated with Tukey's post-hoc test. EC50 values were calculated using log-logistic models in the *drc* package in R (Ritz et al., 2015). Measured concentrations from experimental vessels were used for all 1-methylnaphthalene calculations, and dilutions of metal concentrations used in exposures were calculated from measured stocks of copper sulphate and zinc sulphate.

$EC50_{min}$ values were to be calculated according to a model described by French-McCay (2002); however, EC50s were not calculable for enough time points to complete the calculations. Dunnett's post-hoc test was used to determine statistically significant differences between treatments and controls for both exposure and recovery periods. Control variability was assessed by calculating a CV between replicate control units. Statistically significant differences between EC50 values for each compound between the three trials were assessed using the *EDcomp* function in R. Significance level for all statistical tests conducted in this study was $p < 0.05$.

3.3 Results

3.3.1 Optimization Trials

3.3.1.1 Media optimization trials

The media optimization trial yielded significantly different results between media types for both *N. frigida* and *P. glacialis* as shown in Figure 3.1. Media types prepared with NSW grew significantly more than those prepared with ASW for F/2 prepared in-house and Harrison's medium. There was no significant difference between seawater types for the commercial formulation of F/2. The greatest growth occurred in F/2 prepared in-house with NSW (mean yield = 88,284 cells/mL) and HESNW (mean yield = 88,162 cells/mL), both of which were statistically significant relative to other media types; however; the insignificant difference in cell yield between these two media types and the relative ease of preparing Harrison's media resulted in HESNW being used as the preferred media type for further testing.

HESAW yielded the most growth for *P. glacialis* (mean yield = 20,492 cells/mL), though this result was not statistically different from HESNW (mean yield = 15,705

cells/mL). There were no statistically significant differences between either of the F/2 preparations regardless of seawater type. As there was a minimal and non-significant difference in growth between *P. glacialis* cultures grown in HESAW or HESNW and *N. frigida* exhibited significantly more growth in NSW, HESNW was used for both species in subsequent testing to facilitate solution preparation.

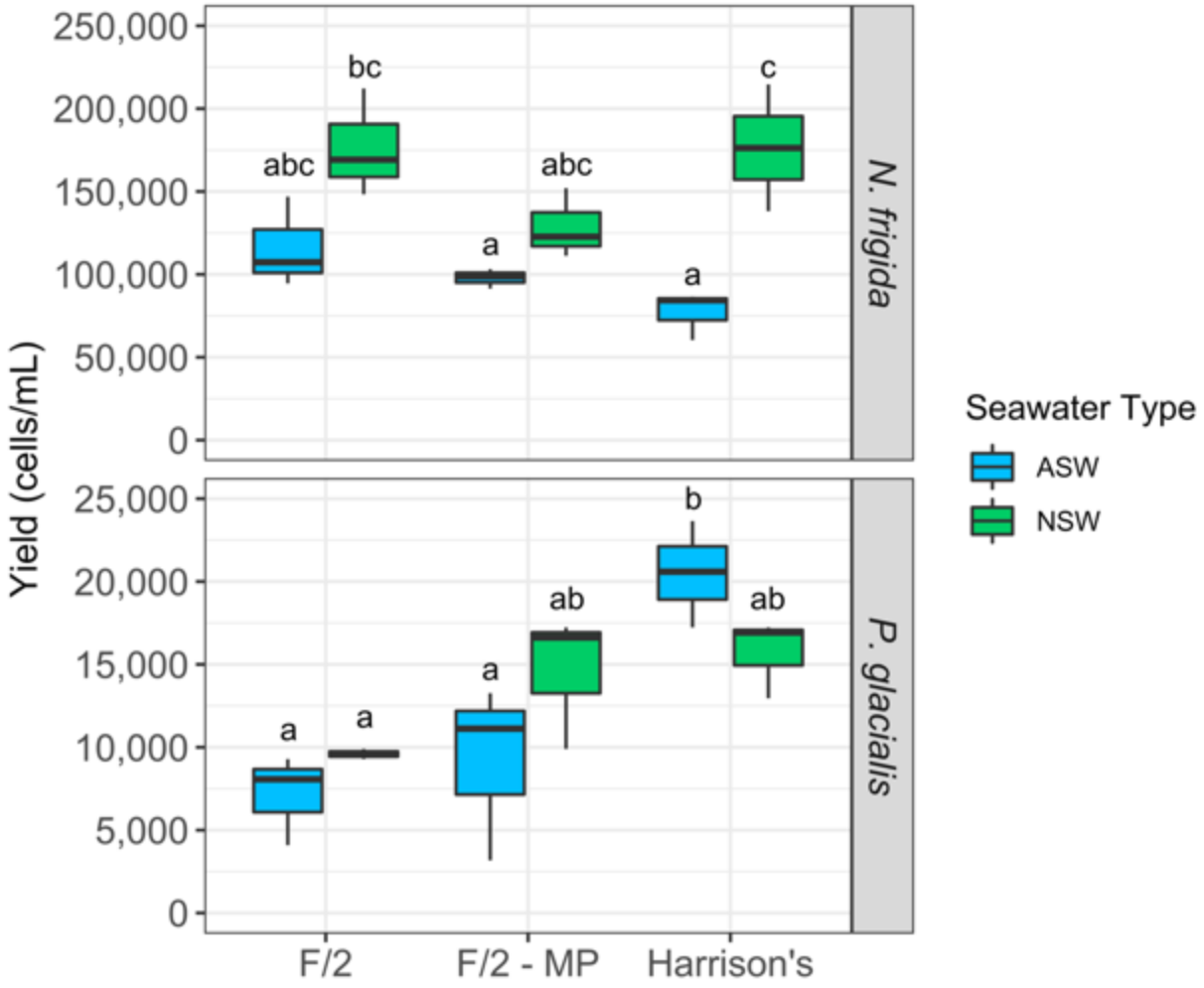

Figure 3.1 Box plots of 14-day growth responses of *Nitzschia frigida and Porosira glacialis* incubated under various media types. F/2 = Guillard's F/2 prepared from scratch; F/2 – MP = Guillard's F/2 prepared from mass packs; Harrison's = Harrison's medium prepared from scratch; ASW = artificial seawater; NSW = natural seawater. Letters indicate statistically significant differences between media types ($p < 0.05$) based on one-way ANOVA followed by Tukey's HSD test.

3.3.1.2 Test vessel optimization trials

The test vessel optimization trial resulted in no statistically significant differences in test vessel type for either species. Microplate data was discounted upon analysis as accurate concentrations could not be quantified from the spectrophotometry values. Although growth in the wells could be observed macroscopically, this did not uniformly translate into higher readings from the instrument. Moving the microplate from a 2°C environment to the microplate reader at room temperature resulted in condensation on the well plate even though the temperature change was brief. This could have resulted in the variability shown in Figures S39-40. For this reason, a LOD was not able to be ascertained for the microplate method with this cold-water species. Figure 3.2 shows the results from the flask and test tube units only for both species. As there were no significant differences between test vessel types for each species, test tubes were used in further testing to minimize space and volume requirements.

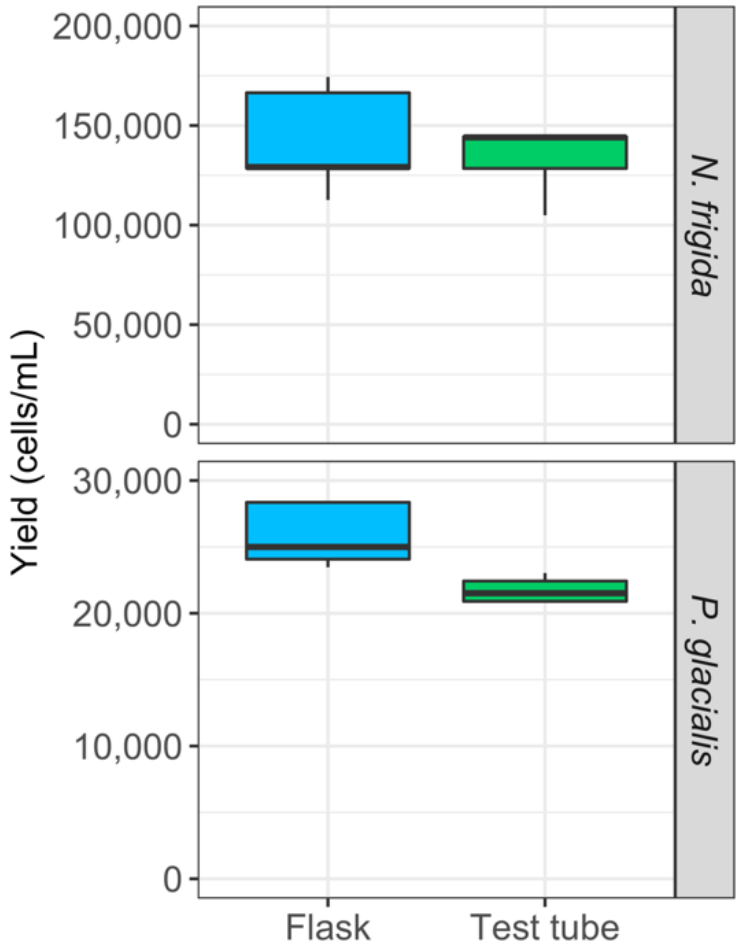

Figure 3.2 Box plots of 14-day growth responses of *Nitzschia frigida* and *Porosira glacialis* incubated in different test vessel types. No statistically significant differences ($p < 0.05$) were identified between test vessel types for each species based on one-way ANOVA followed by Tukey's HSD test.

3.3.1.3 Temperature and light trials

As demonstrated in Figure 3.3, *N. frigida* did not exhibit any growth at 10 or 14°C under any of the light regimes. At 6°C, growth was significantly inhibited at a light intensity of 120 µmol photons m^{-2} s^{-1} PAR relative to 60 and 90 µmol photons m^{-2} s^{-1} PAR at the same temperature. At 2°C, 90 µmol photons m^{-2} s^{-1} PAR resulted in significantly greater growth relative to 60 and 120 µmol photons m^{-2} s^{-1} PAR at the same temperature. Overall, there was no significant difference in growth between 2°C at a light intensity of 90 µmol photons m^{-2} s^{-1} PAR, 6°C at a light intensity of 60 µmol photons m^{-2} s^{-1} PAR, and 6°C at a light intensity of 90 µmol photons m^{-2} s^{-1} PAR. To mirror an

Arctic environment more closely, 2°C was selected as the temperature and 90 µmol photons m^{-2} s^{-1} PAR as the light intensity to be used in further testing.

P. glacialis exhibited minimal growth in this trial, similarly to the media and test vessel optimization trials. Though there were statistically significant differences in growth at various temperatures and light intensities as shown in Figure 3.3, none of the combinations yielded sufficient growth to continue considering *P. glacialis* as a viable option for definitive testing. *N. frigida* was selected as the Arctic species to use for range-finding and definitive toxicity tests.

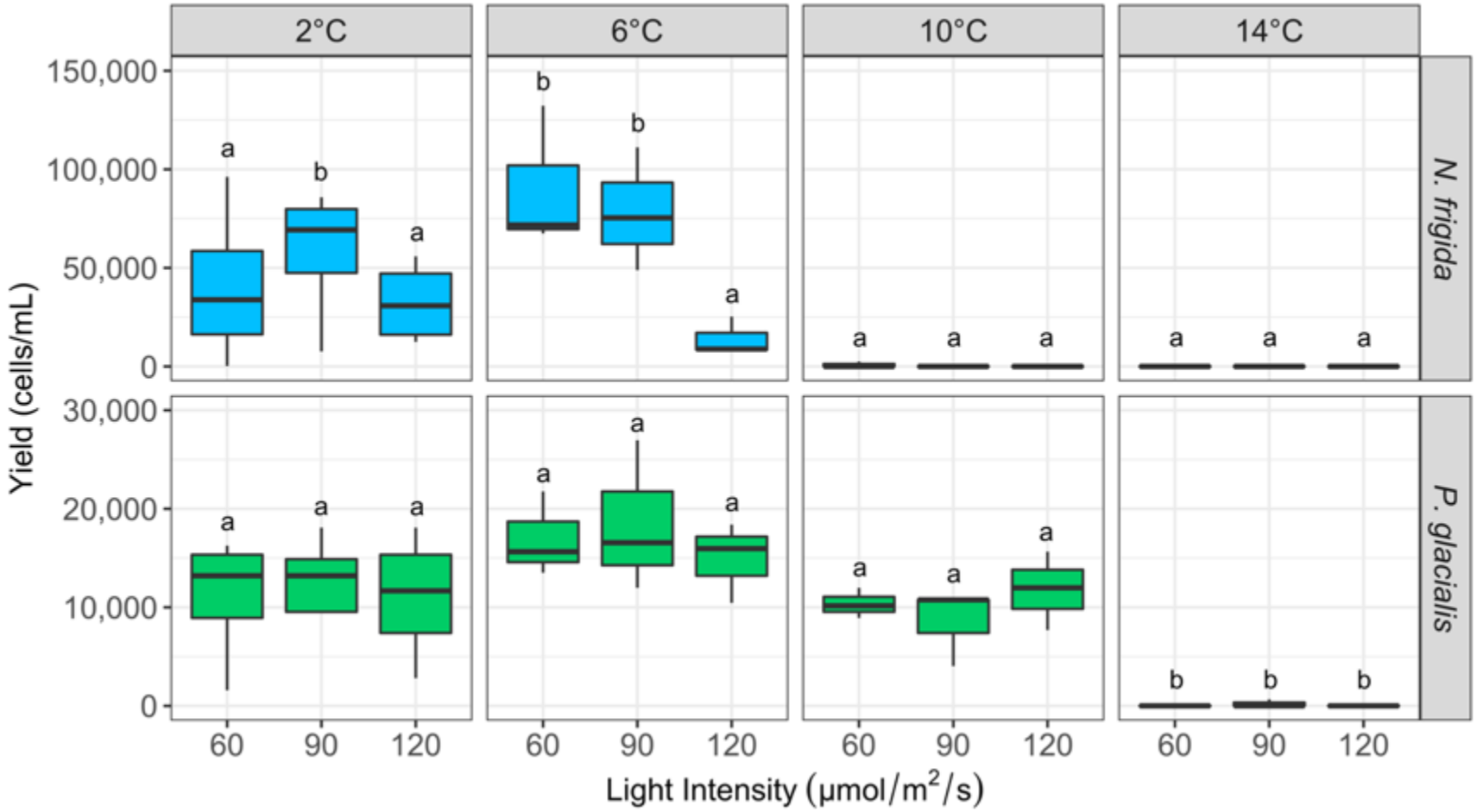

Figure 3.3 Box plots of 14-day growth responses of *Nitzschia frigida* and *Porosira glacialis* incubated under various temperatures and light intensities. Letters indicate statistically significant differences between test conditions for each species ($p < 0.05$) based on one-way ANOVA followed by Tukey's HSD test.

3.3.2 Toxicity Tests

3.3.2.1 Range-finding tests

The results of the range-finding tests conducted with *N. frigida* are presented in Figure 3.4. Cell yield at concentrations of 1000 µg copper/L and above were statistically different from controls; thus, definitive test concentrations between 450 and 2278 µg/L were selected for this compound. Cell yield at 1000 µg zinc/L was similar to controls, and 2000 µg/L resulted in almost total inhibition of growth; thus, definitive concentrations between 450 and 2278 µg/L were selected for this compound as well. Finally, cell yield at concentrations of 100 µg/L 1-methylnaphthalene and above were statistically different from controls and resulted in almost total inhibition of growth; thus, definitive concentrations to capture the appropriate range of effect between 30 and 2820 µg/L were selected for this compound. Range-finding test results for *S. costatum* can be found in Figure S41.

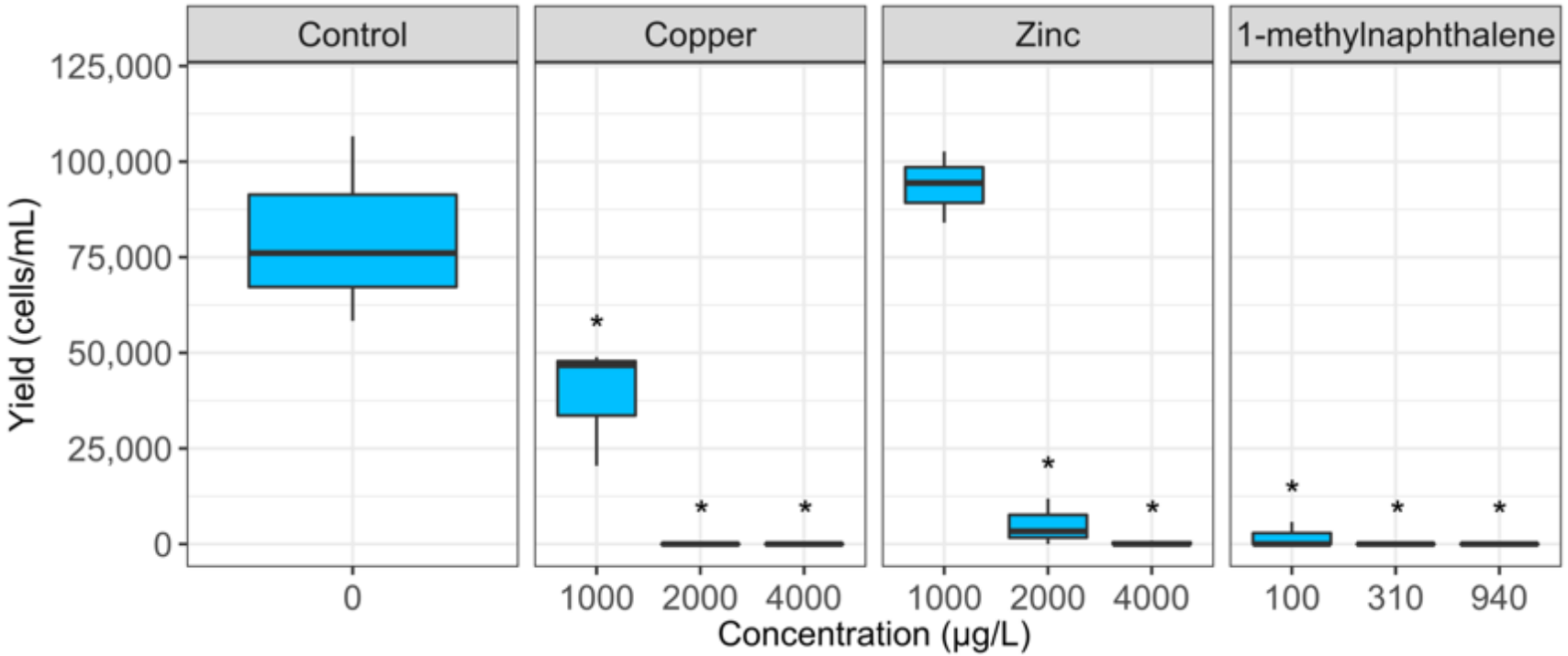

Figure 3.4 Boxplots of 14-day growth responses of *Nitzschia frigida* exposed to copper, zinc, and 1-methylnaphthalene. Asterisks denote statistically significant differences ($p < 0.05$) from control based on one-way ANOVA followed by Dunnett's post-hoc test.

3.3.2.2 Definitive tests

The standard protocol presented in this paper for *N. frigida* has an exposure duration of 14 days. As previously described, additional sampling was performed on days 8, 10, 12, and 16 in an effort to produce multiple EC50 values from which incipient lethal levels could be calculated to facilitate a sensitivity comparison between species; however, insufficient growth in some of the earlier sampling points for *N. frigida* resulted in incalculable or inaccurate EC50 calculations. This is discussed in further detail in Section 4.1.4. Results presented below relate to the assessment of consistency of the newly developed protocol for the 14-day time point for *N. frigida* only. EC50 values from tests conducted with *S. costatum* following are included below to compare the relative sensitivity of the Arctic and temperate species to the three tested compounds at the standard protocol end of exposure only. Results from 96-hour standard tests with *S. costatum* from which these values were calculated can be found in the Supplemental Information of this article (Figure S42), along with calculated values for all other sampling points with *S. costatum.* For yield, the endpoint from which EC50 values were calculated, 222 of 1526 data points (*n* = 98 from *N. frigida* tests and *n* = 124 from *S. costatum* tests) were removed due to negative values obtained as described in the Methods.

3.3.2.2.1 Environmental conditions

Pre-exposure solution temperatures for *N. frigida* ranged from 0.6 to 3.2ºC across all treatments and trials. Salinity of test solutions ranged from 30.6 to 32.4 psu, DO saturation ranged from 90.6 to 99.8% saturation, and pH values ranged from 8.44 to 8.69. Pre-exposure solution temperatures for *S. costatum* ranged from 17.0 to 19.8ºC across all treatments and trials, salinity ranged from 31.6 to 33.6 psu, DO saturation ranged from 96.5 to 101.1% saturation, and pH values ranged from 8.28 to 8.63.

Pre-exposure measurements in pH surrogate units for *N. frigida* ranged from 8.43 to 8.81 and post-exposure measurements ranged from 7.50 to 8.57. For *S. costatum,* pre-exposure values were 8.30 to 8.68 and post-exposure were 7.63 to 8.99.

A summary of pre-exposure water quality, as well as pre-exposure and post-exposure pH values from pH surrogate units, can be found in Table 3.2. Individual values can be found in Table S7.

Table 3.2 Summary of measured test conditions from three trials conducted in the present study. Values are presented as the minimum, maximum, and average ± standard deviation of all measurements.

		Skeletonema costatum	*Nitzschia frigida*
Pre-Exposure Water Quality (test solutions)			
Temp. (°C)	Min.	17	0.6
	Max.	19.8	3.2
	Average ± SD	18.7 ± 1.0	2.1 ± 0.5
Salinity (psu)	Min.	31.6	30.6
	Max.	33.6	32.4
	Average ± SD	32.6 ± 0.7	31.4 ± .6
O_2 (% sat.)	Min.	96.5	90.6
	Min.	101.1	99.8
	Max.	97.8 ± 1.2	95.3 ± 2.8
pH	Min.	8.28	0.44
	Max.	8.63	8.69
	Average ± SD	8.45 ± 0.12	8.57 ± 0.07
Light (µmol photons m^{-2} s^{-1} PAR)	Min.	82	81
	Max.	106[a]	115
	Average ± SD	92[a] ± 7	90 ± 7
Pre-Exposure pH (surrogate units)			
	Min.	8.301	8.431
	Max.	8.667	8.809
	Average ± SD	8.470 ± 0.111	8.574 ± 0.098
Post-Exposure pH (surrogate units)			
	Min.	7.626	7.495
	Max.	8.994	8.573
	Average pH ± SD	8.417 ± 0.388	8.247 ± 0.226

[a] One anomalous value of 211 µmol photons m^{-2} s^{-1} PAR not included.

Over the course of all trials, light intensity ranged from 81 to 115 µmol photons m^{-2} s^{-1} PAR for *N. frigida* and 82 to 106 µmol photons m^{-2} s^{-1} PAR for *S. costatum*, with the exception of a single anomalous measurement of 211 µmol photons m^{-2} s^{-1} PAR where the lighting was adjusted immediately upon observation. Summary statistics of the illumination measurements for both species can be found in Table 3.2.

3.3.2.2.2 Inoculum controls

ICs demonstrated more variability for *N. frigida* than for *S. costatum* as shown in Figure S32. The coefficients of variation (CV) among the five ICs for *N. frigida* were 28%, 18%, and 26% in trials 1, 2, and 3, respectively, while the CV among the five ICs for *S. costatum* were 15%, 13%, and 24%.

3.3.2.2.3 Control performance

Control performance was the first measure analyzed to assess consistency within and between the three definitive tests that were performed. A benchmark for acceptable control variability in many standard algal testing protocols is achieving a CV of less than 20% between control replicates (e.g., ASTM, 2012; ECCC, 2007; NIWA, 1998). As demonstrated in Figure 3.5, this criterion was met in two of the three definitive trials conducted with *N. frigida.* In the first trial, the mean yield for controls was 63,176 cells/mL with a CV of 15.4%. The second trial yielded less mean control growth at 41,077 cells/mL but also had acceptable variability with a CV of 17.7%. The third trial performed similarly to the second in terms of mean growth with 46,475 cells/mL but was slightly more variable with a CV of 24.6%. O-ring controls in trial 2 grew slightly less than regular controls and were more variable (mean cell yield = 32,229 cells/mL, CV = 49%). O-ring controls in trial 3 grew significantly more than regular controls and were slightly less variable (mean cell yield = 91,065 cells/mL, CV = 23%, Figure 3.6).

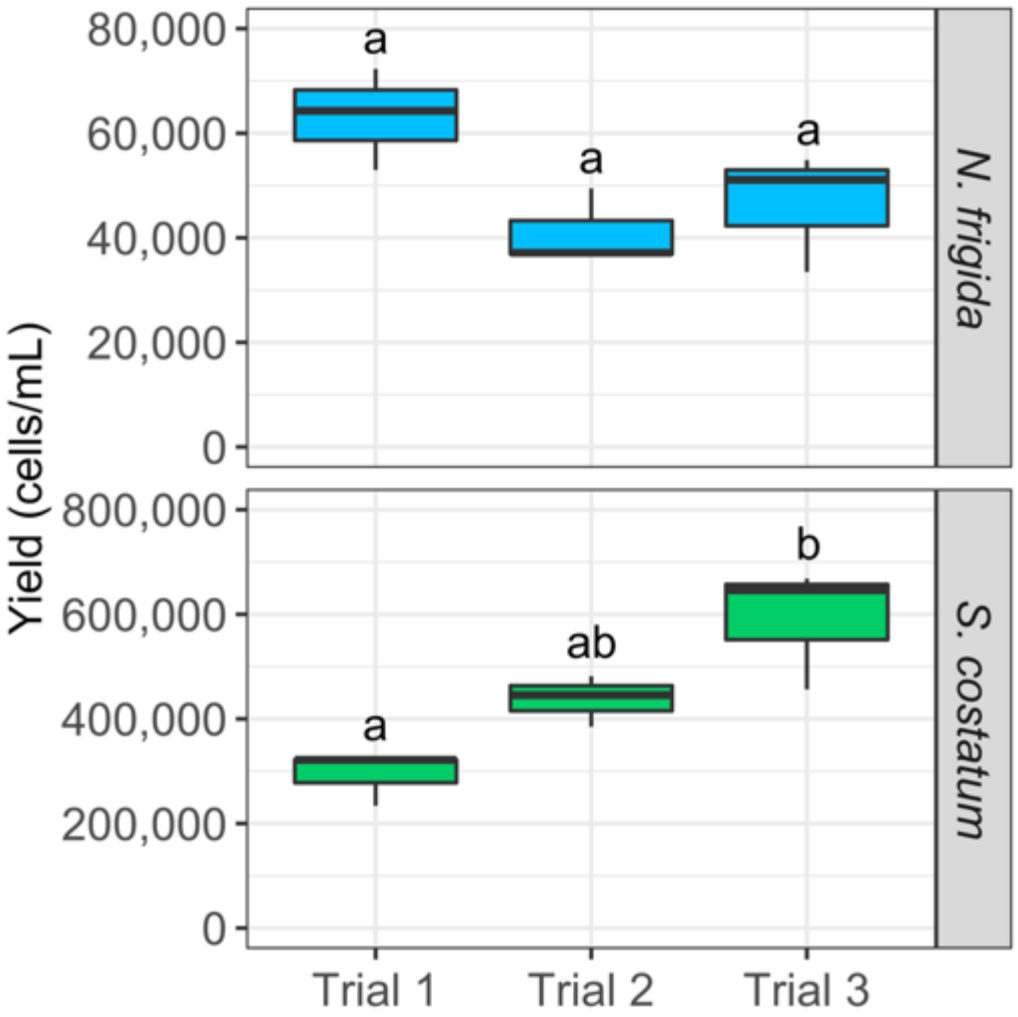

Figure 3.5 Performance of control units after 14-days (*Nitzschia frigida*) or 96 hours (*Skeletonema costatum*) of incubation in three definitive toxicity tests. Letters indicate statistically significant differences between trials for each species ($p < 0.05$) based on one-way ANOVA followed by Tukey's HSD test.

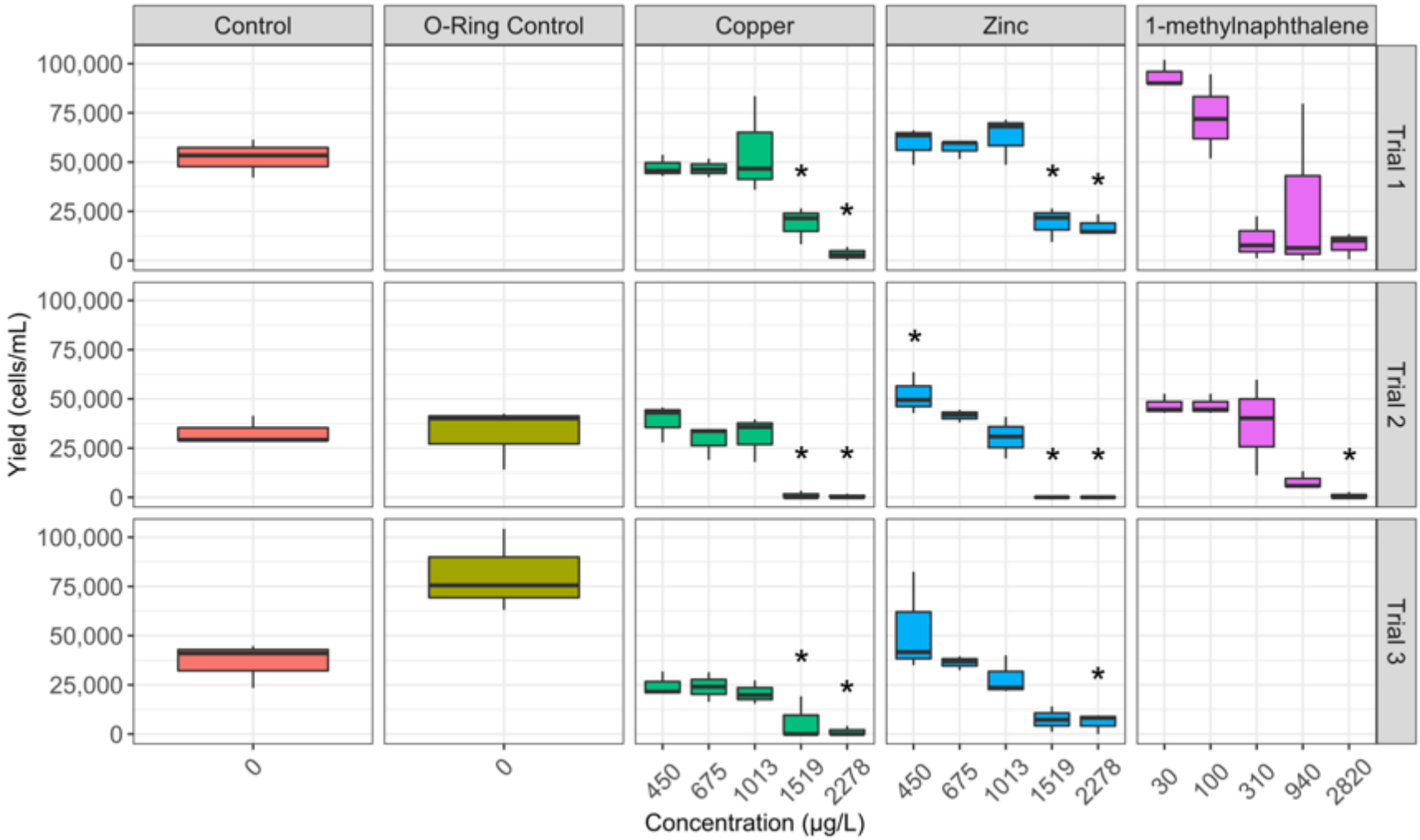

Figure 3.6 Box plots of 14-day growth responses of *Nitzschia frigida* exposed to copper, zinc, and 1-methylnaphthalene. Asterisks denote statistically significant differences ($p < 0.05$) from controls based on one-way ANOVA followed by Dunnett's post-hoc test where data were normally distributed with equal variances, and based on Kruskal-Wallis rank sum test followed by Wilcoxon signed-rank test where data were not normally distributed.

3.3.2.2.4 Standard protocol endpoint results (14-day EC50)

Growth responses from three trials for *N. frigida* after 14 days of exposure are presented in Figure 3.6. Cultures responded in a concentration-dependent manner for all compounds in all three trials except 1-methylnaphthalene in trial 3. Concentration-response curves were generated for each compound and test in order to calculate 14-day EC50 values based on cell yield. Figure 3.7 depicts concentration-response curves for this species for all three trials; however, this relationship was not always statistically significant. The 14-day EC50 values for copper were significant in all three tests; for zinc in all three tests; and for 1-methylnaphthalene in one of three tests. Only statistically significant EC50 values were used to gauge consistency of response across trials and to compare relative sensitivities with *S. costatum*. The 14-day EC50 values for copper ± standard error were 1071 ± 421 µg/L, 1204 ± 238 µg/L, and 1484 ± 437 µg/L. None of the calculated values for this compound were statistically different from one another, demonstrating a consistent response for this compound. The three responses to zinc were also consistent with one another with EC50 values of 1075 ± 178 µg/L, 1076 ± 364 µg/L, and 1386 ± 576 µg/L. The single statistically significant EC50 value for 1-methylnaphthalene was 328 ± 103 µg/L. *p*-values for comparisons between 14-day EC50 values from the three trials can be found in Table 3.3.

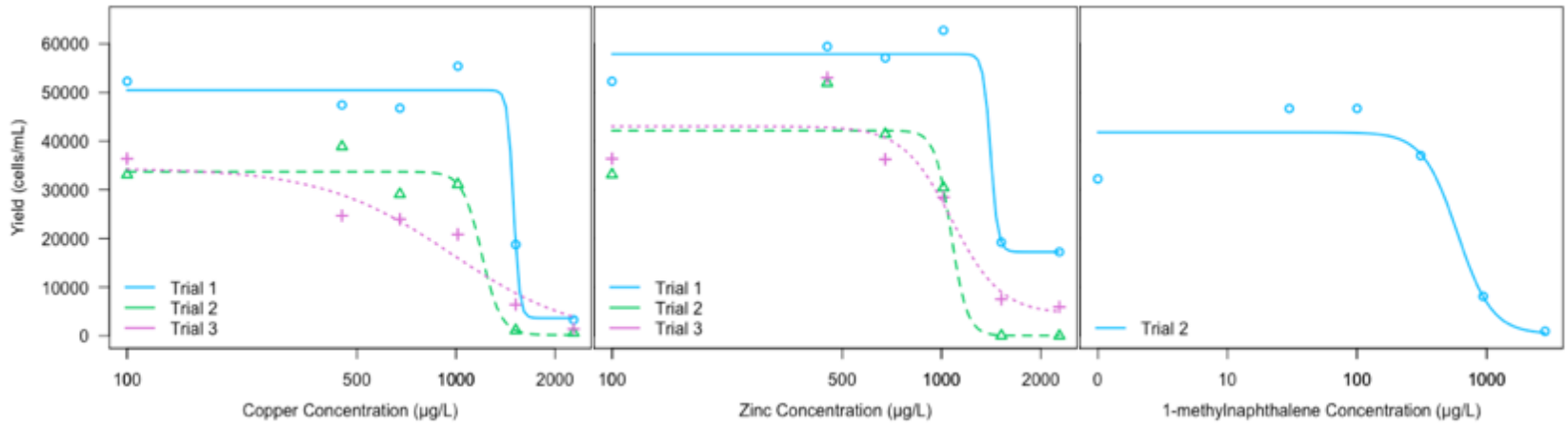

Figure 3.7 Concentration-response curves generated using the *drc* package in R from 14-day toxicity tests with *Nitzschia frigida*. Curves for 1-methylnaphthalene in Trials 1 and 3 were not statistically significant ($p < 0.05$) and thus are not shown.

Table 3.3 Statistical significance of differences between EC50 values calculated from three discrete 14-day exposures of *Nitzschia frigida* to copper, zinc, and 1-methylnaphthalene. *p*-values were calculated using the *EDcomp* function in R.

Test Substance	Trial ID	*p*-value
Copper	1 - 2	NC[a]
	1 - 3	NC[a]
	2 - 3	0.689
Zinc	1 - 2	0.869
	1 - 3	0.854
	2 - 3	0.993
1-methylnaphthalene	1 - 2	0.752
	1 - 3	0.549
	2 - 3	NC[a]

[a] NC = Not calculable

3.3.2.2.5 EC50 values over time

EC50 values for all time points and trials can be found in Table S8. Statistically significant EC50 values at each time point are presented in Table 3.4. Figure 3.8 shows the statistically significant EC50 values over time and indicates consistency of response for each particular sampling day. For each time point where an EC50 value was accurately calculated, there were no statistically significant differences between trials with the exception of 1-methylnaphthalene in Trial 3.

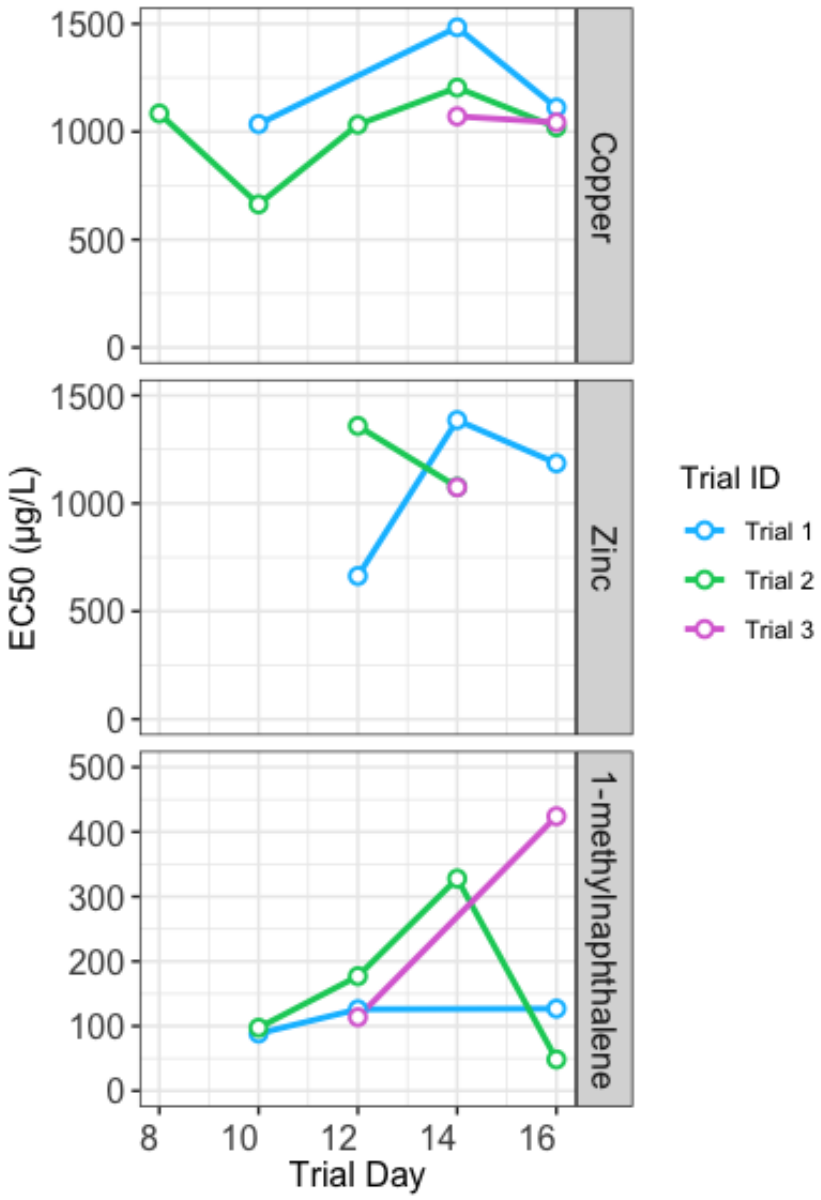

Figure 3.8 Statistically significant ($p < 0.05$) 14-day EC50 values from tests conducted with *Nitzschia frigida* and various test substances on days 8, 10, 12, 14, and 16 of exposure. EC50 values and statistical significance of the relationships were calculated using the *drc* package in R.

Table 3.4 EC50 values based on cell yield from various time points in three discrete exposures of *Nitzschia frigida* to copper, zinc, and 1-methylnaphthalene. Values generated from concentration-response curves that were not statistically significant ($p < 0.05$) are not presented. Results were calculated according to a four-parameter log-logistic model using the *drc* package in R.

Species	Test Substance	Trial Day	EC50 (yield) (µg/L)	Standard Error	Lower Limit (95% CI)	Upper Limit (95% CI)	*p*-value
N. frigida	Copper	8	1084.5	304.5	431.5	1737.5	3.00×10^{-3}
		10	1035.3	173.2	663.8	1406.8	3.39×10^{-5}
			663.4	246.3	135.1	1191.8	0.02
		12	1032.8	101.1	815.9	1249.6	7.18×10^{-8}
		14	1483.8	437.4	545.7	2421.8	4.00×10^{-3}
			1204.3	237.5	694.8	1713.7	1.00×10^{-4}
			1071.1	420.9	168.4	1973.8	0.02
		16	1111.9	94.4	909.4	1314.5	1.19×10^{-8}
			1020.0	43.8	926.2	1113.9	1.34×10^{-12}
			1043.2	232.3	545.0	1541.3	5.00×10^{-4}
	1-methylnaphthalene	10	88.5	30.3	23.5	153.5	0.01
			97.4	20.3	53.8	141.0	3.00×10^{-4}
		12	125.9	28.4	64.9	187.0	6.00×10^{-4}
			177.1	20.5	133.2	220.9	5.39×10^{-7}
			113.7	10.7	90.7	136.7	4.49×10^{-8}
		14	327.9	103.3	106.3	549.4	7.00×10^{-3}
		16	127.0	49.3	21.2	232.7	0.02
			48.5	7.2	33.0	63.9	9.79×10^{-6}
			424.1	104.1	200.8	647.5	1.00×10^{-3}
	Zinc	12	663.3	81.8	488.0	838.8	1.17×10^{-6}
			1358.9	550.2	178.7	2539.0	0.03
		14	1385.7	575.6	151.3	2620.2	0.03
			1076.3	364.4	294.6	1857.9	0.01
			1074.5	178.4	691.9	1457.2	3.00×10^{-5}
		16	1185.7	347.0	441.6	1929.9	4.00×10^{-3}
S. costatum	Copper	4	2129.0	334.5	1411.7	2846.4	1.75×10^{-5}
			2801.0	237.0	2292.8	3309.3	1.14×10^{-8}
			2412.8	199.1	1985.7	2839.9	8.26×10^{-9}
		5	2156.7	140.8	1854.7	2458.7	3.86×10^{-10}
			2339.9	108.4	2107.3	2572.5	3.83×10^{-12}
			2478.1	141.5	2174.5	2781.7	6.48×10^{-11}

Species	Test Substance	Trial Day	EC50 (yield) (µg/L)	Standard Error	Lower Limit (95% CI)	Upper Limit (95% CI)	*p*-value
		6	2285.4	115.9	2036.9	2533.9	1.30×10^{-11}
			2258.1	180.5	1871.1	2645.1	5.45×10^{-9}
		7	2193.0	33.1	2122.1	2263.9	2.20×10^{-16}
			2424.2	187.4	2022.3	2826.1	3.55×10^{-9}
			2634.3	36.8	2555.3	2713.3	2.20×10^{-16}
	1-methylnaphthalene	4	236.7	102.6	16.5	456.8	0.04
			137.4	6.5	123.3	151.3	5.23×10^{-12}
		5	91.6	21.8	44.9	138.2	9.00×10^{-4}
			38.9	17.2	2.1	75.8	0.04
		6	474.4	190.0	66.9	881.8	0.03
			268.5	43.2	175.8	361.1	2.25×10^{-05}
			47.9	6.9	33.0	62.8	7.12×10^{-6}
		7	212.1	37.0	132.7	291.5	5.20×10^{-5}
			182.7	2.5	177.2	188.1	2.20×10^{-16}
	Zinc	4	2255.3	95.2	2051.0	2459.6	1.08×10^{-12}
			2501.8	46.0	2403.1	2600.5	2.20×10^{-16}
			2123.5	34.2	2050.1	2196.9	2.20×10^{-16}
		5	2511.6	216.4	2047.4	2975.8	1.44×10^{-8}
			2215.7	122.4	1953.1	2478.2	4.14×10^{-11}
			2386.5	51.6	2275.8	2497.0	2.20×10^{-16}
		6	2639.9	366.9	1852.9	3426.9	4.60×10^{-6}
			2221.4	146.1	1908.0	2534.9	4.27×10^{-10}
			2384.2	100.9	2167.7	2600.6	1.11×10^{-12}
		7	2059.0	115.7	1810.8	2307.3	5.22×10^{-11}
			2550.2	19.5	2507.2	2593.2	2.20×10^{-16}
			2701.2	42.2	2610.7	2791.7	2.20×10^{-16}

3.3.2.2.6 Relative sensitivity of *N. frigida* and *S. costatum*

For copper, a significantly higher EC50 value was observed for *S. costatum* (96-hours) than for *N. frigida* (14 days) in all three tests (Table 3.4). Zinc exposure yielded similar results with all three EC50 values for *S. costatum* being significantly higher than the three values for *N. frigida.* The single EC50 value for *N. frigida* exposed to 1-methylnaphthalene was significantly higher than either of the two tests with *S. costatum*.

3.3.2.3 Recovery

Due to minimal growth in the units exposed to the two highest concentrations of copper and zinc (1519 and 2278 µg/L) and three highest concentrations of 1-methylnaphthalene (310, 940, and 2820 µg/L) throughout the exposure, cells from these units were not transferred to recovery. These units did not contain a sufficient number of cells with which to inoculate recovery units and therefore could not be assessed for recovery potential.

Recovery data indicated that cultures were able to recover post-exposure after being transferred to clean media. Among all three trials, there was no statistically significant concentration-dependent response in recovery growth for any of the three compounds (Figure 3.9). Trials 1 and 3 demonstrated no statistically significant differences in cell yield among any of the previously exposed treatments compared to the control. In Trial 2, cells previously exposed to the 675 µg/L concentration of zinc exhibited significantly more cell growth in the recovery period than controls; however, this significant relationship was not concentration-dependent as this was the sole concentration where this occurred. These responses were echoed in the recovery testing with *S. costatum* (Figure S43). Despite the ability of exposed cultures to recover relative to control recovery, cell yield was lower overall following the recovery period compared to the exposure period. Mean control cell yield for Trial 1 recovery was 41,879 cells/mL compared to 63,176 cells/mL after exposure incubated for the same duration; Trial 2 recovery was 31,176 cells/mL compared to 41,077 cells/mL after exposure; and trial 3 recovery was 42,879 cells/mL compared to 46,475 cells/mL after exposure.

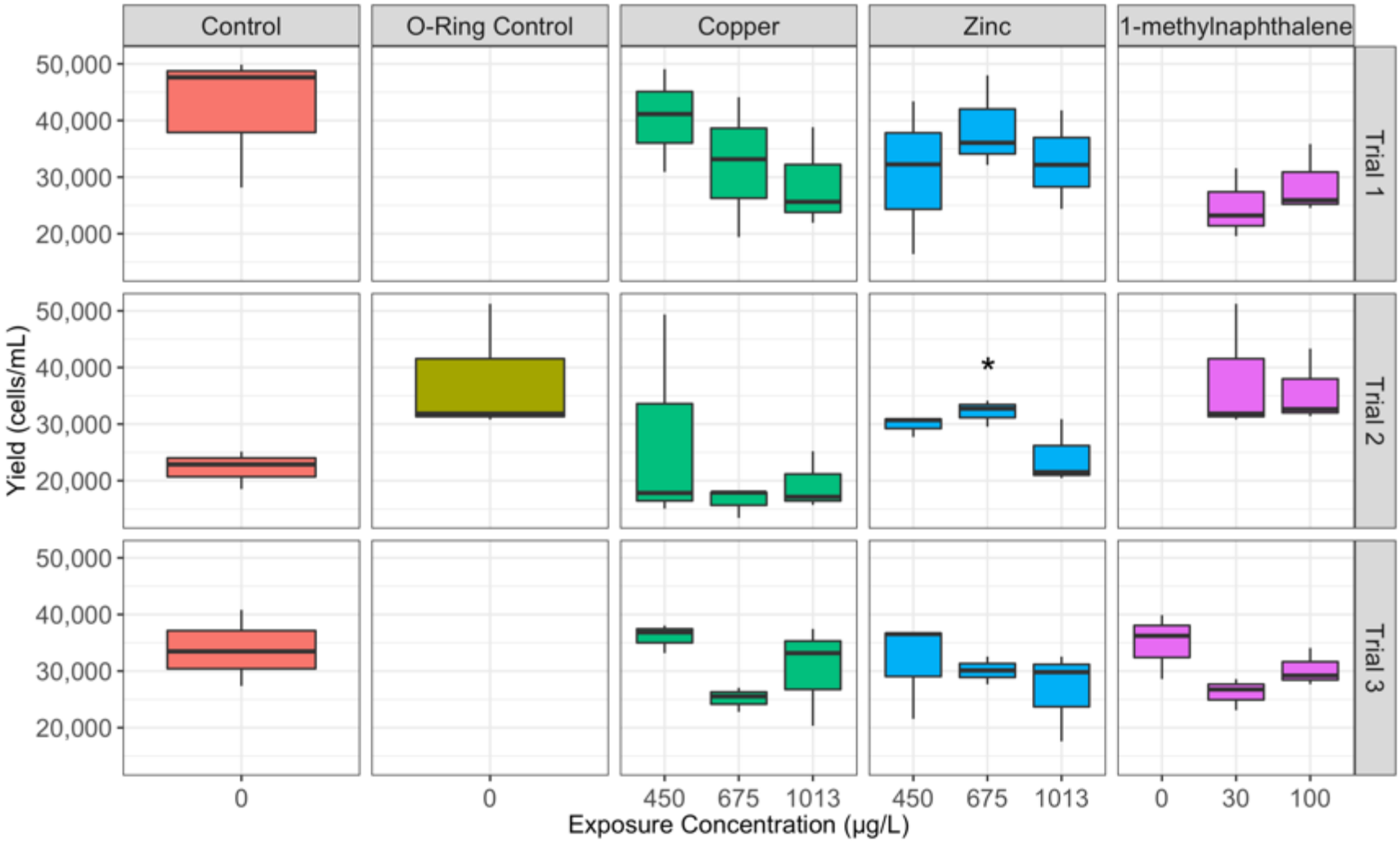

Figure 3.9 Box plots of growth responses of *Nitzschia frigida* after 14 days of exposure to various test substances followed by 14 days of recovery in clean media. Asterisks denote statistically significant differences ($p < 0.05$) from controls based on one-way ANOVA followed by Dunnett's post-hoc test.

3.4 Discussion

The overall objectives of this study were met by successfully optimizing culturing conditions for an Arctic diatom for use in toxicity testing, designing toxicity tests that proved to be consistent and repeatable based on a number of criteria, and comparing the relative sensitivities to toxicants and recovery potential of the new species with a standard diatom. These results produced evidence that *N. frigida* can be more sensitive to select compounds than that of the temperate diatom *S. costatum*, but there are several caveats that may render their sensitivity differences negligible. The lower EC50 values for this species and these compounds were still within an order of magnitude of the responses from *S. costatum*, even though the Arctic species test had a 14-day exposure duration compared to 96 hours for the temperate species. Although this could not be avoided due to the nature of their respective environments, it renders the results more difficult to compare. Additionally, risk assessors often apply an uncertainty factor to risk and hazard calculations to account for unknown variables and factors in the environment. This, combined with the relatively close EC50 values obtained from each species despite very different exposure durations, present a scenario where responses from the temperate diatom *S. costatum* may be sufficiently protective of the Arctic diatom *N. frigida* for first tier risk assessments or hazard quotient calculations. In some cases, it may be more economical and feasible to encourage the reporting of results from temperate diatom tests due to this minor difference in sensitivity; however, the environmental realism captured in both the exposure conditions and test species in the protocol developed here may be more valuable than more easily achieved data that cannot mimic these conditions. For these reasons, use of *N. frigida* is recommended as an excellent candidate for a standard test species representative of an Arctic environment.

3.4.1 Consistency of the Protocol

The newly developed protocol for an Arctic diatom generally produced consistent results across three separate trials; however, there are several areas that could be

targeted to improve the overall repeatability of the protocol and ensure consistent results across space and time. These are discussed below.

3.4.1.1 Control performance

ICs had higher CVs within and across trials for *N. frigida* and this may be attributable to the colonial nature of these cultures. Although both *N. frigida* and *S. costatum* are chain-forming, *N. frigida* forms arborescent colonies that aggregate whereas *S. costatum* forms single-file chains. When viewed under the microscope, it was more common to observe single cells of *S. costatum* not attached to others than it was for *N. frigida*.

The CV between replicates for controls was < 20% in two of the three trials conducted and was only slightly above 20% in the third test (24.6%). Given the colonial nature of *N. frigida* cells and the tendency of cells to adhere to vessel walls instead of remaining suspended in the water column, some amount of variability was to be expected, even if only as a result of sampling. Even after thorough mixing of each test vessel before removing a sample to obtain the algal concentration in the unit, microscopic colonies of cells continue to be present which limits the degree of homogeneity that is possible to achieve. Another approach that could be taken with respect to reducing this variability in future modifications to the protocol would be the addition of continuous mixing, either by use of a magnetic stir bar, bubbling, or orbital shaking. This was not attempted in the present study as previous experience with continuous shaking of colonial diatoms in culture flasks resulted in aggregation of cells, but this approach was not tested with *N. frigida* specifically.

Cell yield in control units among all three tests was sufficient to discern differences between controls and treatments at the end of 14 days of exposure; however, growth was variable between tests. Where two of the tests yielded more similar results (Trial 2 at 41,077 cells/mL and Trial 3 at 46,475 cells/mL) than the first (Trial 1 at 63,176 cells/mL), it can be expected that control growth would more consistently fall in the 40,000 to 50,000 cells/mL range. More tests would need to be run to strengthen this argument, but even with growth in this range a concentration-response could be observed from control growth to the highest treatments.

The increased growth in O-ring control units for Trial 3 relative to regular controls lines up with the behaviour and physiology of ice algae in the Arctic environment as surface-associated pennate diatoms (Lund-Hansen et al., 2020). These organisms thrive attached to a substrate in nature, and the results from this study indicate that the addition of a substrate to EUs may increase growth rates in exposures by appeasing their tendency to adhere to their surroundings. However, the average fold increase for regular controls (i.e., not O-ring controls) across all three trials was approximately 5-fold the average density of the respective ICs, which was sufficient for our purposes in this protocol and thus was not further explored. Greater control growth could increase the sensitivity of the test in future iterations of the protocol and may result in the ability to detect more subtle effects between controls and treatments than are presented here.

3.4.1.2 Response to toxicants

Consistency of responses across tests in the present study could only be measured for copper and zinc as an EC50 value for 1-methylnaphthalene was only calculable for one of the three trials. Growth inhibition as a result of metal exposure is highly variable depending on the algal species studied and the specific metal tested. Franklin et al. (2009) found that several species exposed to copper had 48 or 72-hour EC50 values of 10 µg/L or less, while Cid et al. (1994) reported values of over 200 µg/L for tests of the same duration. EC50 values of 30 – 1210 µg/L were observed in tests conducted by Rojìčková-Padrtová and Maršálek (1999). Many factors influence metal toxicity to algae and may present reasons for this range of values (e.g., initial cell density of EUs studied in Debelius et al. [2009] as well as the release of metal-binding exudates by the cells themselves and accumulation kinetics in the algal culture [Paquet et al., 2015]). EC50 values for copper at the end of the exposures conducted herein were in line with values reported in the literature. In addition, they were not statistically different in any of the three trials conducted and thus presented consistency of response under these exposure conditions. Zinc exposure also yielded consistent results between trials with EC50 values of 1075, 1076, and 1386 µg/L. Zinc toxicity to other marine algal species has been reported at EC50 concentrations of 10 – 2310 µg/L for tests of similar duration (Rojìčková-Padrtová and Maršálek, 1999).

The significantly higher EC50 values for *S. costatum* in copper and zinc tests indicate that *N. frigida* may be more sensitive to these compounds. Moreover, the exposure conditions in this protocol are more similar to what Arctic diatoms would experience in nature, reducing uncertainty around how water temperature may impact toxicity. These data may add to the weight of evidence (WOE) in identifying hazardous concentrations of metals in Arctic marine environments for ERAs conducted there. Future testing could include the use of metals with different characteristics, such as arsenic or cadmium, which would further allow the data to be compared against regulatory guidelines (e.g., Canadian Environmental Quality Guidelines by CCME).

Mechanisms of copper toxicity specifically to algal cells have been reported in the literature but are largely dependent on species. For example, Franklin et al. (2009) found that some freshwater species first responded to the compound with changes in cell membrane permeability or local esterase activity. Marine diatoms were reported as exhibiting possibly depolarization in the cells as a result of copper exposure. Disruptions in photosynthesis, respiration, and cell division have also been noted as a result of copper binding to the cell membrane (Johnson et al., 2007; Stauber and Florence, 1987). These mechanisms of action are important to note in the context of recovery potential as they help us better understand how populations may react in the long-term, after contaminants are no longer present in their ecosystem.

Consistency of response between trials could not be assessed for 1-methylnaphthalene as only a single statistically significant EC50 value could be calculated; however, at 328 µg/L it is significantly lower than the copper and zinc EC50s. From the data available for comparison, it can be deduced that *N. frigida* is significantly more sensitive to 1-methylnaphthalene than it is to either copper or zinc. *N. frigida* and *S. costatum* both exhibited sensitivities to PAHs that are in line with what has been observed in the literature for other algal species (e.g., *P. subcapitata* EC50 values for growth rate of 509 and 1045 µg/L in Bragin et al., 2016).

The toxicity of 1-methylnaphthalene appears to be greater in *S. costatum* compared to *N. frigida* based on evidence presented in this study. Oil spill research in general is highly variable due to its inherent complexity; it is made up of thousands of compounds, and there are a wide variety of approaches taken in the literature for

reporting toxicity (e.g., based on TPHs, sum of PAHs, or mass of crude oil). This makes it difficult to compare toxicity of different oils or compounds found in oil across organisms and environments. However, the present study provides evidence that the standard test species may be more sensitive to oil components than the Arctic diatom selected for this work, indicating that *S. costatum* may be an appropriate test species in this specific context. However, further testing with *N. frigida* against other single compounds found in oil such as pyrene or whole crude or heavy fuel oils would assist in understanding its sensitivity to this class of substances, which is critical in developing a deeper understanding of how populations of these organisms may react in the event of an oil or fuel spill.

As a final note, this laboratory culture may have responded differently to toxicants than what may be observed from the exposure of a freshly-isolated culture. Physiological and biological changes have been observed in phytoplankton populations cultured in the laboratory for long periods of time relative to their counterparts in nature, which may influence how they respond to contaminants. For example, deformations and mutations, slower growth rates, and adaptations to growth media have all been observed in laboratory algal cultures when compared to field cultures (Lakeman et al., 2009). The culture of *N. frigida* used in this study has been cultured in the laboratory for several years and exhibits differences in size and shape relative to when it was first isolated. In nature, *N. frigida* cells are long and straight; however, the cells in this culture are shorter and curved. Availability and accessibility of Arctic diatoms in general is limited in practice, and thus it is difficult to obtain cultures that have not had the opportunity to adapt to laboratory conditions. It would be beneficial to conduct further research assessing whether freshly-isolated cultures mirror the responses outlined within this thesis.

3.4.2 Recovery

The lack of a concentration-dependent pattern of statistically significant differences between controls and treatments in the recovery portion of the definitive tests indicates that the recovery observed did not correlate with exposure treatment, and previously exposed cultures were able to fully recover relative to controls for all

compounds tested. Less growth was observed overall in recovery units, but since growth did not correlate with treatment it can be assumed that this was a result of handling in the transfer to recovery process and not as a result of exposure. From these data it can be deduced that *N. frigida* cells were in fact able to recover from exposure to copper, zinc, and 1-methylnaphthalene after being rinsed and re-suspended in clean media.

The implications of recovery potential are important in the context of ERA. There are innumerable scenarios of environmental contamination that can result in anything from long-term, chronic exposures (e.g., oil from a spill being trapped in the under-ice environment) to short-term pulses (e.g., wastewater effluent release), which can be subsequently diluted by oceanic or wind-driven currents. Additionally, ice algae may be exposed to the more stable compounds of a particular contaminant for long periods of time as the more volatile compounds dissipate into the atmosphere. For example, 1-methylnaphthalene exposure would likely be the result of an oil spill where short-term concentrations of this specific compound would be high, but would taper off over time. Other experiments have demonstrated that PAH compounds also have the capability to move through the ice-associated ecosystem by migrating through the sea-ice layer itself, providing another avenue by which exposure to microorganisms may eventually be lessened (Boccadoro et al., 2018). This would be determined by a variety of factors, such as proximity to the source of contamination, water and air currents, and other environmental conditions. An algal population that is only temporarily inhibited has fewer and less severe implications for communities and ecosystems that are the primary focus of ERA.

Full recovery of algal cultures after exposure, such as that presented herein, has been observed throughout the literature. The primary test substances for which recovery potential has been assessed have been herbicides. Vallotton et al. (2008) assessed the time-to-recovery of a freshwater alga after exposure to *S*-metolachlor and determined that previously exposed cultures took more time to restart cell division after being transferred to clean media than controls. Experiments by Baxter et al. (2014) demonstrated a similar recovery response as was observed with *N. frigida* in the present study (i.e., no statistically significant difference in growth after a recovery period

between controls and previously exposed populations) after exposure of atrazine to a freshwater alga. The full recovery observed with *N. frigida* and *S. costatum* herein indicates that short-term (i.e., 14-days and 96 hours respectively) metal and PAH exposure under natural conditions will inhibit growth while the contaminant is present, but populations will likely be restored in the absence of contamination. As growth endpoints were only measured at the end of the recovery period, the data presented here is not indicative of how long that may take; therefore, future studies to address this (i.e., similar to Vallotton et al., 2008) would be beneficial in garnering further understanding of this process.

3.4.3 Incipient lethal level

The original aim of sampling from EUs at four additional time points outside of the identified exposure end (14 days) was to collect data that could be used to calculate incipient lethal levels of each compound to this Arctic species (e.g., Philibert et al., 2021). Unfortunately, the earlier time points (days 8, 10, and 12) did not consistently yield enough growth in the controls or treatments from which accurate EC50 values could be calculated. The relatively slower growth of Arctic species compared to temperate species resulted in a much longer exposure duration for *N. frigida* than *S. costatum*, rendering the results more difficult to compare. To obtain an appropriate number of EC50 values to be used in $EC50_{min}$ calculations, the test would likely need to be extended to 20 – 22 days to ensure sufficient growth is achieved in test vessels. Other endpoints, such as average specific growth rate, could be further examined to address differences in exposure duration. Regardless of this element, the environmental realism of the exposure conditions in the protocol designed in this study may bear sufficient weight for certain ERAs to render its use regardless of the availability and sensitivity of a temperate species test.

3.5 Conclusion

Overall, this study met our objectives to develop a new protocol that could be used as a tool for risk assessors in conducting ERAs in Arctic environments. We were

able to optimize culture conditions for an Arctic diatom such that a toxicity test could be run within a reasonable time frame (Table 3.5), perform the test and demonstrate repeatability and consistency of results, and compare the relative sensitivity of the Arctic species with a standard, temperate species. This work contributes to ecotoxicology and ERA by providing a new approach to add to the WOE in Arctic ERAs and helping to help reduce uncertainty regarding environmental effects.

Table 3.5 Summary of test conditions for the 14-day *Nitzschia frigida* toxicity testing protocol developed in the present study.

Species	*Nitzschia frigida*
Test Type	Static
Exposure Duration	14 days
Vessel	25-mL borosilicate glass test tubes
Medium	Harrison's with natural seawater
Temperature (ºC)	2 ± 2
Salinity (psu)	30 - 32
pH	8.2 - 8.6
Light Quality	Cool-white fluorescent
Light Intensity (µmol photons m^{-2} s^{-1} PAR)	90 ± 10%
Photoperiod	Continuous
Shaking	Manually once daily
Inoculum Density (cells/mL)	10,000
Control CV (%)	≥ 20
Control Growth Criteria	≥ 5-fold increase
Endpoints	EC50 cell yield

3.6 References

American Society for Testing and Materials (ASTM). (2012). Standard guide for conducting static toxicity tests with microalgae. E1218-04.

Arctic Monitoring and Assessment Programme (AMAP). (2007). *Oil and Gas Activities in the Arctic – Effects and Potential Effects*. AMAP, Oslo. Retrieved from https://www.amap.no/documents/download/1016/inline.

Artamonova, E. Y., Svenning, J. B., Vasskog, T., Hansen, E. and Eilertsen, H. C. (2017a). Analysis of phospholipids and neutral lipids in three common northern cold water diatoms: *Coscinodiscus concinnus, Porosira glacialis,* and *Chaetoceros socialis,* by ultra-high performance liquid chromatography-mass spectrometry. *Journal of Applied Phycology, 29*(3), 1241-1249.

Artamonova, E. Y., Vasskog, T., Eiliertsen, H. C., Tsang, Y. F. (2017b). Lipid content and fatty acid composition of *Porosira glacialis* and *Attheya longicornis* in response to carbon dioxide (CO_2) aeration. *PLoS One, 12*(5), e017703-e017703.

Aumack, C. F. and Juhl, A. R. (2015). Light and nutrient effects on the settling characteristics of the sea ice diatom *Nitzschia frigida. Limnology and Oceanography, 60*, 765-776.

Aune, M., Aschan, M. M., Greenacre, M., Dolgov, A. V., Fossheim, M., Primicerio, R. (2018). Functional roles and redundancy of demersal Barents Sea fish: Ecological implications of environmental change. *PLoS One, 13*(11), e0207451.

Baxter, L., Brain, R., Rodriguez-Gil, J. L., Hosmer, A., Solomon, K., Hanson, M. (2014). Response of the green alga *Oophila* sp., a salamander endosymbiont, to a PSII-inhibitor under laboratory conditions. *Environmental Toxicology and Chemistry, 33*(8), 1858-1864.

Bejarano, A. C., Gardiner, W. W., Barron, M. G., Word, J. Q. (2017). Relative sensitivity of Arctic species to physically and chemically dispersed oil determined from three hydrocarbon measures of aquatic toxicity. *Marine Pollution Bulletin, 122*(1-2), 316-322.

Boccadoro, C., Krolicka, A., Receveur, J., Aeppli, C., Le Floch, S. (2018). Microbial community response and migration of petroleum compounds during a sea-ice oil spill experiment in Svalbard. *Marine Environmental Research, 142,* 214-233.

Bradstreet, M. S. W. and Cross, W. E. (1982). Trophic relationships at High Arctic ice edges. *Arctic, 35*(1), 1-12.

Bragin, G. E., Parkerton, T. F., Redman, A. D., Letinksi, D. J., Butler, J. D., Paumen, M. L., Sutherland, C. A., Knarr, T. M., Comber, M., den Haan, K. (2016). Chronic toxicity of selected polycyclic aromatic hydrocarbons to algae and crustaceans using passive dosing. *Environmental Toxicology and Chemistry, 35*(12), 2948-2957.

Butler, J. D., Parkerton, T. F., Redman, A. D., Letinski, D. J., Cooper, K. R. (2016). Assessing aromatic-hydrocarbon toxicity to fish early life stages using passive-dosing methods and target-lipid and chemical-activity models. *Environmental Science and Technology, 50*, 8305-8315.

Camus, L., Brooks, S., Geraudie, P., Hjorth, M., Nahrgang, J., Olsen, G. H., Smit, M. G. D. (2015). Comparison of produced water toxicity to Arctic and temperate species. *Ecotoxicology and Environmental Safety, 113*, 248-258.

Chapman, P. M., Riddle, M. J. (2005). Toxic effects of contaminants in polar marine environments. *Environmental Science and Technology,* 201A-207A.

Chaves-Barquero, L. G., Luong, K. H., Mundy, C. J., Knapp, C. W., Hanson, M. L., and Wong, C. S. (2016). The release of wastewater contaminants in the Arctic: A case study from Cambridge Bay, Nunavut, Canada. *Environmental Pollution*, *218*, 542–550.

Cid, A., Fidalogo, P., Herrero, C., Abaide, J. (1994). Different responses of the marine diatom *Phaeodactylum tricornutum* to copper toxicity. *Microbiologia, 10*, 263-272.

Cota, G. F., Legendre, L., Gosselin, M., Ingram, R. G. (1991). Ecology of bottom ice algae: I. Environmental controls and variability. *Journal of Marine Systems, 2*, 257-277.

Croft, M. T., Lawrence, A. D., Raux-Deery, E., Warren, M. J., Smith, A. G. (2005). Algae acquire vitamin B_{12} through a symbiotic relationship with bacteria. *Nature, 438*, 90-93.

Darnis, G., Robert, D., Pomerleau, C., Link, H., Archambault, P., Nelson, R. J., Geoffroy, M., Tremblay, J.-E., Lovejoy, C., Ferguson, S. H., Hunt, B. P. V., Fortier, L. (2012). Current state and trends in Canadian Arctic marine ecosystems: II. Heterotrophic food web, pelagic-benthic coupling, and biodiversity. *Climatic Change, 115*(1), 179-205.

Debelius, B., Forja, J. M., DelValls, Á., Lubián, L. M. (2009). Toxicity and bioaccumulation of copper and lead in five marine microalgae. *Ecotoxicology and Environmental Safety, 72*(5), 1503-1513.

Environment and Climate Change Canada (ECCC). (2007). Biological Test Method: Growth Inhibition Test Using a Freshwater Alga. EPS/1/RM/25 Second Edition.

Environment and Climate Change Canada (ECCC). (2013, July 11). *Wildlife and landscape science research topics: Biological test method development.* Government of Canada. https://www.canada.ca/en/environment-climate-change/services/wildlife-research-landscape-science/research-topics/biological-test-method-development.html.

Fernández-Méndez, M., Wenzhöfer, F., Peeken, I., Sørensen, H. L., Glud, R. N., and Boetius, A. (2014). Composition, buoyancy regulation and fate of ice algal aggregates in the Central Arctic Ocean. *PLoS ONE*, *9*(9).

Food and Agriculture Organization of the United Nations. (2004). *Hatchery culture of bivalves: A practical manual.* FAO Fisheries Technical Paper 471. Food and Agriculture Organization of the United Nations.

Franklin, N. M., Stauber, J. L., Lim, R. P. (2009). Development of flow cytometry-based algal bioassays for assessing toxicity of copper in natural waters. *Environmental Toxicology and Chemistry, 20*(1), 160-170.

French-McCay, D. P. (2002). Development and application of an oil toxicity and exposure model, OilToxEx. *Environmental Toxicology and Chemistry, 21*, 2080-2094.

Gosselin, M., Levasseur, M., and Wheeler, P. A. (1997). Deep sea research Part II: Topical studies in oceanography - New measurements of phytoplankton and ice algal production in the Arctic Ocean. *Deep Sea Research*, *44*(8).

Guillard, R. R. L. (1975). Culture of phytoplankton for feeding marine invertebrates. In W. L. Smith and M. H. Chanley (Eds.), *Culture of marine invertebrate animals* (pp. 26-60). Plenum Press.

Guillard, R. R. L. (2005). Purification methods for microalgae. In R. Andersen (Ed.), *Algal culturing techniques,* (pp. 117-132). Academic Press.

Guillard, R. R. L. and Ryther, J. H. (1962). Studies of marine planktonic diatoms. I. *Cyclotella nana hustedt,* and *Detonula confervacea* (Cleve) Gran. *Canadian Journal of Microbiology, 8*(2) 229-239.

Gunnarsdóttir, R., Jenssen, P. D., Jensen, E. P., Villumsen, A., Kallenborn, R. (2013). A review of wastewater handling in the Arctic with special reference to pharmaceuticals and personal care products (PPCPs) and microbial pollution. *Ecological Engineering, 50,* 76-85.

Ha, S.-Y., La, H. S., Min, J.-O., Chung, K.-H., Kang, S.-H., Shin, K.-H. (2014). Photoprotective function of mycosporine-like amino acids in a bipolar diatom (*Porosira glacialis*): Evidence from ultraviolet radiation and stable isotope probing. *Diatom Research, 29*(4), 399-409.

Hansen, B., Altin, D., Bonaunet, K., Øverjordet, I. (2014). Acute toxicity of eight oil spill response chemicals to temperate, boreal, and Arctic species. *Journal of Toxicology and Environmental Health, Part A, 77*, 495-505.

Health Canada. (2018, June 28). *Environmental Toxicology.* Government of Canada. https://www.canada.ca/en/health-canada/services/consumer-product-safety/pesticides-pest-management/public/protecting-your-health-environment/pesticide-registration-process/reviews/environmental-toxicology.html.

Hermanson, M. H., Jackson, D. S., Johnson, G. W. (2010). History of Inuit community excretion of polychlorinated biphenyls recorded in sewage lake sediments. *Chemosphere, 78*(11), 1322-1328.

International Organization for Standardization. (2016). *Water quality – Marine algal growth inhibition test with Skeletonema* sp. *and* Phaeodactylum tricornutum (ISO Standard No. 10253:2016).

Johnson, H. L., Stauber, J. L., Adams, M., Jolley, D. F. (2007). Copper and zinc tolerance of two tropical microalgae after copper acclimation. *Environmental Toxicology and Chemistry, 22*, 234-244.

Johnson, K. (2008). National Inuit position paper regarding the CCME Canada-Wide strategy for the management of municipal wastewater effluent and Environment Canada's proposed regulatory framework for wastewater. Ottawa: Inuit Tapiriit Kanatami.

Kallenborn, R., Brorström-Lundén, E., Reiersen, L. O., Wilson, S. (2018). Pharmaceuticals and personal care products (PPCPs) in Arctic environments: Indicator contaminants for assessing local and remote anthropogenic sources in a pristine ecosystem in change. *Environmental Science and Pollution Research, 25*, 33001-33013.

Kan, Y. and Pan, J. (2010). A one-shot solution to bacterial and fungal contamination in the green alga *Chlamydomonas reinhardtii* culture by using an antibiotic cocktail. *Journal of Phycology, 46*(6), 1356-1358.

Kohlbach, D., Graeve, M., Lange, B. A., David, C., Peeken, I., Flores, H. (2016). The importance of ice algae-produced carbon in the central Arctic Ocean ecosystem: Food web relationships revealed by lipid and stable isotope analysis. *Limnology and Oceanography, 61*(6), 2027-2044.

Kohlbach, D., Schaafsma, F. L., Graeve, M., Lebreton, B., Lange, B. A. David, C., Vorkamp, M., Hauke, F. (2017). Strong linkage of polar cod (*Boreogadus saida*) to sea ice algae-produced carbon: Evidence from stomach content, fatty acid and stable isotope analyses. *Progress in Oceanography, 152,* 62-74.

Kooistra, W. H. C. F., Sarno, D., Balzano, S., Gu, H., Andersen, R. A., Zingone, A. (2008). Global diversity and biogeography of *Skeletonema* species (Bacillariophyta). *Protist, 159*(2), 177-193.

Krumhansl, K. A., Krkosek, W. H., Greenwood, M., Ragush, C., Schmidt, J., Grant, J., Barrell, J., Lu, L., Lam, B., Gagnon, G. A., Jamieson, R. C. (2014). Assessment of Arctic community wastewater impacts on marine benthic invertebrates. *Environmental Science and Technology, 49*(2), 760-766.

Lakeman, M. B., von Dassow, P., Cattolico, R. A. (2009). The strain concept in phytoplankton ecology. *Harmful Algae, 8*, 746-758.

Letcher, R. J., Bustnes, J. O., Dietz, R., Jenssen, B. M., Jørgensen, E. H., Sonne, C., Verreault, J., Vijayan, M. M., Gabrielsen, G. W. (2010). Exposure and effects assessment of persistent organohalogen contaminants in Arctic wildlife and fish. *Science of the Total Environment, 408*(15), 2995-3043.

Lund-Hansen, L. C., Søgaard, D. H., Sorrell, B. K., Gradinger, R., Meiners, K. M. (2020). Spring, Summer, and Melting Sea Ice. In J. Ford (Ed.), *Arctic Sea Ice Ecology: Seasonal Dynamics in Algal and Bacterial Productivity* (pp. 61-101). Springer Polar Sciences, Switzerland.

McMinn, A., Pankowski, A., Delfatti, T. (2005). Effect of hyperoxia on the growth and photosynthesis of polar sea ice microalgae. *Journal of Phycology, 41*(4), 732-741.

Meng, F., Huang, W., Liu, D., Zhao, Y., Huang, W., Lei, Z., Zhang, Z. (2020). Application of aerobic granules-continuous flow reactor for saline wastewater treatment: Granular stability, lipid production and symbiotic relationship between bacteria and algae. *Bioresource Technology, 295*, 122291.

Michel, C., Legendre, L., Ingram, R. G., Gosselin, M., and Levasseur, M. (1996). Carbon budget of sea-ice algae in spring: Evidence of a significant transfer to zooplankton grazers. *Journal of Geophysical Research: Oceans, 101*(C8), 18345-18360.

Michel, C., Nielsen, T. G., Nozais, C., and Gosselin, M. (2002). Significance of sedimentation and grazing by ice micro- and meiofauna for carbon cycling in annual sea ice (northern Baffin Bay). *Aquatic Microbial Ecology, 30*(1), 57-68.

Mundy, C .J. and Meiners, K. (2021). Chapter 10: Sea Ice Ecology. In D. N. Thomas (Ed.). *Arctic Ecology* (textbook). Wiley-Blackwell, Oxford, UK.

National Institute of Water and Atmospheric Research (NIWA). (1998). Appendix 1 (Marine algae [*Dunaliella tertiolecta*] chronic toxicity test protocol) to Standard Methods for Whole Effluent Toxicity Testing: Development and Application. Report no. MFE80205. NIWA Report for the Ministry for the Environment, Wellington, New Zealand.

Nuttall, M., Berkes, F., Forbes, B., Kofinas, G., lassova, T., Wenzel, G. (2005). Hunting, herding, fishing, and gathering: Indigenous peoples and renewable resource use in the Arctic. *Arctic Climate Impact Assessment,* 649-690.

Olsen, G. H., Smit, M. G. D., Carroll, J., Jaeger, I., Smith, T., Camus, L. (2011). Arctic versus temperate comparison of risk assessment metrics for 2-methylnaphthalene. *Marine Environmental Research, 72,* 179-187.

Organization for Economic Cooperation and Development. (2011). OECD Guidelines for the Testing of Chemicals.

Paquet, N., Lavoie, M., Maloney, F., Duval, J. F. L., Campbell, P. G. C., Fortin, C. (2015). Cadmium accumulation and toxicity in the unicellular alga *Pseudokirchneriella subcapitata*: Influence of metal-binding exudates and exposure time. *Environmental Toxicology and Chemistry, 34*(7), 1524-1532.

Philibert, D., Parkerton, T., Marteinson, S., de Jourdan, B. (2021). Assessing the toxicity of individual aromatic compounds and mixtures to American lobster (*Homarus americanus*) larvae using a passive dosing system. *Environmental Toxicology and Chemistry, 40*(5), 1379-1388.

Poulin, M., Underwood, G. J. C., and Michel, C. (2014). Sub-ice colonial *Melosira arctica* in Arctic first-year ice. *Diatom Research*, *29*(2), 213–221.

R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Renegar, D. A., Turner, N. R., Riegl, B. M., Dodge, R. E., Knap, A. H., Schuler, P. A. (2017). Acute and subacute toxicity of the polycyclic aromatic hydrocarbon 1-methylnaphthalene to the shallow-water coral *Poritesdivaricata*: Application of a novel exposure protocol. *Environmental Toxicology and Chemistry, 36*, 212-219.

Riedel, A., Michel, C., Poulin, M., and Lessard, S. (2003). Taxonomy and abundance of microalgae and protists at a first-year sea ice station near Resolute Bay, Nunavut, spring to early summer 2001. Canadian Data Report of Hydrography and Ocean Sciences 159.

Ritz, C., Baty, F., Streibig, J. C., Gerhard, D. (2015). Dose-response analysis using R. *PLoS ONE, 10*(12), e0146021.

Rojìčková-Padrtová, R. and Maršálek, B. (1999). Selection and sensitivity comparisons of algal species for toxicity testing. *Chemosphere, 38*(14), 3329-3338.

Rózańska, M., Gosselin, M., Poulin, M., Wiktor, J. M., and Michel, C. (2009). Influence of environmental factors on the development of bottom ice protist communities during the winter-spring transition. *Marine Ecology Progress Series*, *386*, 43–59.

Simis, S. G. H., Huot, Y., Babin, M., Seppälä, J., Metsamaa, L. (2012). Optimization of variable fluorescence measurements of phytoplankton communities with cyanobacteria. *Photosynthesis Research, 112*(1), 13-30.

Sison-Mangus, M. P., Jiang, S., Tran, K. N., Kudela, R. M. (2014). Host-specific adaptation governs the interaction of the marine diatom, *Pseudo-nitzschia* and their microbiota. *Multidisciplinary Journal of Microbial Ecology, 8*(1), 63-76.

Søreide, J. E., Leu, E., Berge, J., Graeve, M., Falk-Petersen, S. (2010). Timing of blooms, algal food quality and *Calanus glacialis* reproduction and growth in a changing Arctic. *Global Change Biology, 16*(11), 3154-3163.

Stauber, J. L. and Florence, T. M. (1987). Mechanism of toxicity of ionic copper and copper complexes to algae. *Marine Biology, 94*, 511-519.

Staveley, J. P. and Smrchek, J. C. (2005). Algal Toxicity Test. In C. Blaise and J.-F. Ferard (Eds.), *Small-Scale Freshwater Toxicity Investigations* (pp. 181-202). Springer, Dordrecht.

Stewart, J. E., Marks, L. J., Wood, C. R., Risser, S. M., Gray, S. (1997). Symbiotic relations between bacteria and the domoic acid producing diatom *Pseudo-nitzschia multiseries* and the capacity of these bacteria for gluconic acid/gluconolactone formation. *Aquatic Microbial Ecology, 12*, 211-221.

Stroski, K. M., Luong, K. H., Challis, J. K., Chaves-Barquero, L. G., Hanson, M. L., Wong, C. S. (2020). Wastewater sources of per- and polyfluorinated alkyl substances (PFAS) and pharmaceuticals in four Canadian Arctic communities. *Science of the Total Environment, 708*, 134494.

Tolvanen, A., Eilu, P., Juutinen, A., Kangas, K., Kivinen, M., Markovaara-Koivisto, M., Naskali, A., Salokannel, V., Tuulentie, S., Similä, J. (2019). Mining in the Arctic environment: A review from ecological, socioeconomic, and legal perspectives. *Journal of Environmental Management, 233,* 832-844.

United States Environmental Protection Agency (US EPA). (1971). Algal Assay Procedure Bottle Test.

United States Environmental Protection Agency (US EPA). (1994, March). *ECO Update: Using Toxicity Tests in Risk Assessment* (9345.0-05I). Office of Solid Waste and Emergency Response. https://www.epa.gov/sites/default/files/2015-11/documents/v2no1.pdf.

United States Environmental Protection Agency (US EPA). (1996). Ecological Effects Test Guidelines OPPTS 850:5400: Algal Toxicity, Tiers I and II.

United States Environmental Protection Agency (US EPA). (2012). *Ecological Effects Test Guidelines* (OCSPP 850.4500: Algal Toxicity).

University of British Columbia (UBC). (2012). *Skeletonema costatum*. In *Phytopedia – The Phytoplankton Encyclopedia Project.* Retrieved from https://www.eoas.ubc.ca/research/phytoplankton/diatoms/centric/skeletonema/s_costatum.html.

Vallotton, N., Moser, D., Eggen, R. I. L., Junghans, M., Chèvre, N. (2008). *S*-metolachlor pulse exposure on the alga *Scenedesmus vacuolatus:* Effects during exposure and the subsequent recovery. *Chemosphere, 73*(3), 395-500.

4 Synthesis

The overall objectives of this thesis were to inform ERA by quantitatively characterizing our current state of knowledge in Arctic ecotoxicology and designing a new tool to be used for risk assessments with Arctic species. As it pertains to the first objective, we found that the methodologies applied in the critical review process allowed us to effectively achieve what we set out to do (Chapter 2). We were able to gather metadata in a way that facilitates comparisons between studies as well as having the ability to identify where gaps exist in the research. The scoring rubric yielded results regarding the reliability of the existing dataset that is available for risk assessors, generating a valuable resource to use in Arctic ERAs. In terms of the second objective, we were successful in optimizing culturing conditions for toxicity testing with an Arctic diatom that has had limited use in a laboratory setting, as well as designing a new protocol for evaluating the impacts of contaminant exposure on this species (Chapter 3). Together, this thesis has highlighted the limitations in our understanding of how Arctic species may be impacted by contaminant exposure, and further, how this unique ecosystem as a whole may respond to the increasing pressure it continues to face as a result of anthropogenic activity. Overall, the research presented herein provides a bank of knowledge and new tools from which risk assessors can draw to assist in furthering our understanding of these issues and lessening the environmental consequences of human activities in the Arctic.

4.1 Critical Review

The relatively few studies that were available for this analysis highlights the need for more well-designed studies with relevant species and endpoints designed specifically for ERAs conducted in an Arctic context. High-quality data that meets these demands have been outwardly identified as lacking for well over a decade (Chapman and Riddle 2003, 2005a, 2005b). More importantly, the data that has been collected is largely unreliable for the purposes of ERA (i.e., only 18% of studies meeting all critical reliability criteria outlined in Chapter 2). There are a variety of reasons why researchers may not be studying Arctic species as frequently as temperate ones. Maintaining Arctic

conditions in a laboratory can sometimes be logistically difficult and expensive. Obtaining cultures of Arctic species in the first place is also less feasible than obtaining temperate ones as fewer laboratories and repositories maintain them. Additionally, there are fewer resources available for learning culturing techniques and experimental methods that are specifically tailored to Arctic exposure conditions. However, it is recommended that more funding be secured from a variety of sources to ensure that the ERA community has access to these types of data although they can be more expensive and time consuming to collect. If we, as scientists and private citizens, wish to protect the Arctic environment to the best of our ability, we no longer have the luxury of avoiding the time and expense of conducting tests with Arctic species. Additionally, we recommend that ecotoxicologists, environmental risk assessors, and researchers in general consult a rubric such as the one presented here when designing studies and using data for any purpose to determine its reliability and relevance for their particular application.

The review process worked well for gathering metadata of Arctic ecotoxicology studies, but hindsight demonstrates several areas that could have been improved to maximize the effectiveness of the method. The initial aim of the review was to target only peer-reviewed articles that had been published in an indexed scientific journal, but we became aware of a bank of Arctic data in the form of abstracts from AMOP just prior to beginning the review. These papers were available at the physical conference location in 2018, so it can be assumed that these data were relatively easy to obtain and readily available for researchers and risk assessors working in this field. For this reason, they were included in the analysis, and the review could have benefitted from exploring other avenues where similar data could be found (e.g., Proceedings from the International Oil Spill Conference [IOSC]). Because the availability of ecotoxicity data was a focus of the review, library databases were primarily searched for papers that were available without requiring contacting authors or sifting through conference proceedings. Only a few dozen papers were retrieved and analyzed this way, which does in fact highlight that data availability is an issue that should be addressed. A separate aspect of the review could have been to consider this lack of availability, and further attempts could have been made to gather more data from less conventional

sources in order to increase the total number of studies assessed against the rubric in spite of their limited availability. We acknowledge that there are more data out there than the 48 studies analysed for reliability here and gathering these for scoring would have provided a more thorough collection to be referenced by risk assessors. The studies for which we have presented reliability scores can hopefully demonstrate how the application of the rubric yields valuable information for use in ERA and risk assessors can use this template for any other data they gather. It is critical that researchers aim to reduce the uncertainty that exists when extrapolating temperate data to Arctic environments. We believe the scoring system presented in the review is a reliable tool for ERA practitioners to tease out data reliability and thereby reduce uncertainty in their assessments of these unique ecosystems.

4.2 Laboratory Tests

It has been hypothesized that *N. frigida* in particular *may* form a mutualistic, commensal, or symbiotic relationship with bacteria found in the same ecosystem. A prerequisite of toxicity testing with algae typically requires that the culture to be used is axenic (i.e., free of other organisms and bacteria that may influence toxicity). The culture used in these laboratory tests was unialgal but not axenic upon receipt from the culture curator; therefore, several attempts were made to purify the culture as described in Chapter 3. The hypothesized relationship between the algae and the bacteria, combined with the difficulty experienced in purifying the culture, resulted in purifying only to the point where visible aggregates of bacteria were no longer microscopically visible. To be considered axenic, the culture would need to be streaked on a solid agar plate and incubated for several weeks to confirm that bacterial colonies would not grow. The fact that we are unsure as to whether the diatom needs this bacteria to thrive limited our ability to conduct the toxicity test exactly as standard methods prescribe (i.e., with an axenic culture); further, it adds uncertainty as to whether the bacteria within the culture are influencing toxicity in any way. A recommendation for future research with this species would be to design an experiment that would determine whether these bacteria are in fact necessary to allow *N. frigida* to thrive or whether they could be

removed from the culture to produce an axenic product that would enable researchers to more closely mirror standard toxicity testing protocols.

The selected temperatures of 2, 6, 10, and 14°C for the optimization trials were based on the hypothesis that increased temperature would lead to increased growth, even with an Arctic species, up until a certain point where growth would decrease. The results, however, indicated that decreasing temperature produced the desired effect. Although the growth achieved with the optimized culturing conditions yielded sufficient growth to discern differences in our toxicity tests, greater growth still would allow more subtle changes to be detected between treatments. *N. frigida* has demonstrated exponential growth at temperatures as low as -6°C that could have yielded more growth had they been trialed in the optimization phase (Aletsee and Jahnke, 1992). This may warrant further investigation should researchers aim to increase the sensitivity of the test. As discussed in Chapter 3, the addition of a substrate could also have yielded greater growth as was evidenced in the O-ring controls. Potential options for assessing growth with a substrate could have been included using an object that is relatively inert and does not have the potential to leach compounds into the exposure media. For example, an item such as a glass microscope slide or a screen would provide a surface onto which cells could adhere without compromising the exposure.

Further to the toxicity test results presented in this thesis, field validations of the results would be beneficial in future studies so that ecosystem-level interactions that may influence toxicity can be observed. The results from this study, as with the results from most laboratory-based ecotoxicology studies, are not meant to demonstrate exactly how this marine diatom would respond to contaminant exposure under natural conditions. It is meant to be a first-tier screening tool to determine whether a risk is posed to the organism under the worst-case scenario, as concentrations do not decline over time as they do in nature and there are few, if any, additional biotic or abiotic factors that may alleviate toxicity. With this in mind, the testing conducted herein determined that in a laboratory setting under Arctic conditions, *N. frigida* may be slightly more sensitive to metals and organic contaminants than the temperate diatom *S. costatum*. Field validating studies could be conducted at the Churchill Marine Observatory in Manitoba, where oil spill related Arctic research is being prioritized

(University of Manitoba, n.d.). For reasons described elsewhere in this thesis, oil and oil-related contaminants are of great concern to Arctic algal communities and are thus highly relevant to study. The inclusion of 1-methylnaphthalene in the tests we conducted provide a set of responses against which field results could be compared at the Observatory if the same compound were to be used.

Comparisons between Arctic and temperate species sensitivities have been presented in the literature for a variety of organism groups (e.g., Hansen et al., 2014; Moore, 2016) and this thesis contributes a relative sensitivity comparison for Arctic and temperate primary producers. Trends in sensitivity were not consistent across contaminant types (i.e., *N. frigida* being more sensitive than *S. costatum* to metals but not 1-methylnaphthalene) which highlights the importance of testing different compound classes with various mechanisms of action (e.g., pharmaceuticals and personal care products found in wastewater). Further differences could be identified by comparing ice algae and pelagic phytoplankton found in the same ecosystem, or testing the same species at different temperatures to establish whether toxicity is a result of inherent sensitivity or of the behaviour of different contaminants in various environmental conditions.

4.3 Summary and Conclusions

In summary, the objectives outlined at the commencement of this research were met in both cases. The data collected in both phases of this thesis will help close the gap in Arctic ecotoxicity testing as it relates to ERA. It is our hope that the critical review rubric will support other researchers in their determinations of high quality, reliable data to be used for environmental decision-making in Canada's northern regions and beyond. Results from the three toxicity tests conducted here are insufficient to draw definitive conclusions as to whether *N .frigida*, or Arctic diatoms in general, are more sensitive than the species used in standard reference toxicity tests; however, they contribute valuable information (i.e., EC50 values) regarding how Arctic diatoms respond to metals and organic contaminants. More importantly, the development and application of standard protocol with an Arctic test species satisfies the call by the

ecotoxicological community for data and resources specifically designed to meet the needs of Arctic ERAs.

As environmental monitoring techniques become more sensitive and are being conducted more frequently, we can become aware of contaminants present in the aquatic environment at exceedingly low levels. This screening gives us the opportunity to act before exposures pose a threat to Arctic biota; however, in order to define the thresholds at which deleterious effects will be observed – or more importantly, will not be observed – it is critical that work in this vein continues, and we strive to reduce the uncertainty around Arctic species' sensitivity to contaminants.

4.4 References

Aletsee, L. and Jahnke, J. (1992). Growth and productivity of the psychrophilic marine diatoms *Thalassiosira antarctica* Comber and *Nitzschia frigida* Grunow in batch cultures at temperatures below the freezing point of sea water. *Polar Biology, 11*, 643-647.

Chapman, P. M. and Riddle, M. J. (2003). Missing and needed: Polar marine ecotoxicology. *Marine Pollution Bulletin, 46*(8), 927-928.

Chapman, P. M. and Riddle, M. J. (2005a). Toxic effects of contaminants in polar marine environments. *Environmental Science and Technology, 39*(9), 200A-207A.

Chapman, P. M., and Riddle, M. J. (2005b). Polar marine toxicology – Future research needs. *Marine Pollution Bulletin*, *50*(9), 905–908.

Hansen, B., Altin, D., Bonaunet, K., Øverjordet, I. (2014). Acute toxicity of eight oil spill response chemicals to temperate, boreal, and Arctic species. *Journal of Toxicology and Environmental Health, Part A, 77*, 495-505.

Moore, D., Sibley, P. K., Solomon, K. R. (2016). *Toxicity of salts, metals, and nitrogenous contaminants to cold-water fish under northern conditions*. (Unpublished doctoral dissertation). University of Manitoba, Winnipeg, MB.

Stock, N. L., Furdui, V. I., Muir, D. C. G., Mabury, S. A. (2007). Perfluoroalkyl contaminants in the Canadian Arctic: Evidence of atmospheric transport and local contamination. *Environmental Science and Technology, 41*(10), 3529-3536.

University of Manitoba. (n.d.). *Churchill Marine Observatory.* Churchill Marine Observatory | Centre for Earth Observation Science | University of Manitoba. Retrieved November 22, 2021, from https://umanitoba.chra/earth-observation-science/facilities-labs-vessels/churchill-marine-observatory.

5 Supplemental Information

LIST OF TABLES

LIST OF FIGURES

5.1 Tables

Table S1 List of studies retrieved from literature searches and reason for exclusion (if applicable).

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
1	Andersen, Ø., Frantzen, M., Rosland, M., Timmerhaus, G., Skugor, A., Krasnov, A.	2015	Effects of crude oil exposure and elevated temperature on the liver transcriptome of polar cod (*Boreogadus saida*)	*Aquatic Toxicology 165,* 9-18.	**Included**	[a]
2	Braune, B. M., Scheuhammer, A. M., Crump, D., Jones, S., Porter, E., and Bond, D.	2012	Toxicity of methylmercury injected into eggs of thick-billed murres and Arctic terns	*Ecotoxicology 21,* 2143-2152.	**Included**	
3	Busdosh, M. and Atlas, R. M.	1977	Toxicity of oil slicks to Arctic amphipods	*Arctic 30,* 85-92.	**Included**	[a]
4	Camus, L., Brooks, S., Geraudie, P., Hjorth, M., Nahrgang, J., Olsen, G. H., Smit, M. G. D.	2015	Comparison of produced water toxicity to Arctic and temperate species	*Ecotoxicology and Environmental Safety 113,* 248-258.	**Included**	
5	Carls, M.G. and Korn, S.	1983	Sensitivity of arctic amphipods and fish to petroleum hydrocarbons	Proceedings of the Tenth Annual Aquatic Toxicity Workshop, Halifax, NS, November 7-10, 1983. Ottawa,	**Included**	

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
				Ontario: Environment Canada.		
6	Desforges, J.-P., Levin, M., Jasperse, L., De Guise, S., Eulaers, I., Letcher, R. J., Acquarone, M., Nordoy, E., Folkow, L. P., Jensen, T. H., Grondahl, C., Bertelsen, M. F., Leger, J. S., Almunia, J., Sonne, C., Dietz, R.	2017	Effects of polar bear and killer whale derived contaminant cocktails on marine mammal immunity	*Environmental Science and Technology 51,* 11431-11439.	**Included**	
7	Dussauze, M., Camus, L., Le Floch, S., Pichavant-Rafini, K., Geraudie P., Coquille, N., Amerand, A., Lemaire, P., Theron, M.	2014	Impact of dispersed fuel oil on cardiac mitochondrial function in polar cod *Boreogadus saida*	*Environmental Science and Pollution Research 21*(24)*,* 13779-13788.	**Included**	
8	Engelhardt, R.	1981	Oil pollution in polar bears: exposure and clinical effects	Proceedings of the Arctic Marine Oil Spill Program Technical	**Included**	[a]

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
				Seminar, Edmonton, AB, June 16-18, 1981. Ottawa, Ontario: Environment Canada.		
9	Engelhardt, R.	1983	Behavioural responses of benthic invertebrates exposed to dispersed crude oil	Proceedings of the Arctic Marine Oil Spill Program Technical Seminar, Edmonton, AB, June 14-16, 1983. Ottawa, Ontario: Environment Canada.	**Included**	[a]
10	Faksness, L.-G., Borseth, J. F., Baussant, T., Hansen, B. H., Altin, D., Tandberg, A. H. S., Ingvarsdottir, A., Aarab, N., Nordtug, T.	2011	The effects of different oil spill cleanup technologies on body burden and biomarkers in Arctic marine organisms – a laboratory study	Proceedings of the Arctic Marine Oil Spill Program Technical Seminar, Banff, AB, October 4-6, 2011. Ottawa, Ontario: Environment Canada.	**Included**	[a]
11	Foy, M. D., Sekarak, A.	1978	Acute lethal toxicity of oil/dispersant mixtures to selected Arctic species	Proceedings of the Arctic Marine Oil Spill Program Technical Seminar,	**Included**	

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
				Edmonton, AB, March 15-17, 1978. Ottawa, Ontario: Environment Canada.		
12	Foy, M. G.	1979	Acute lethal toxicity of Prudhoe Bay Crude Oil and Corexit 9527 to Arctic marine invertebrates and fish from Frobisher Bay, N.W.T.	Proceedings of the Arctic Marine Oil Spill Program Technical Seminar, Edmonton, AB, March 7-9, 1979. Ottawa, Ontario: Environment Canada.	**Included**	
13	Frantzen, M., Falk-Petersen, I.-B., Nahrgang, J., Smith, T. J., Olsen, G. H., Hangstad, T. A., Camus, L.	2012	Toxicity of crude oil and pyrene to the embryos of beach spawning capelin (*Mallotus villosus*)	*Aquatic Toxicology 108,* 42-52.	**Included**	
14	Frouin, H., Loseto, L. L. et al.	2012	Mercury toxicity in beluga whale lymphocytes: limited effects of selenium protection	*Aquatic Toxicology 109,* 185-193.	**Included**	
15	Gardiner, W. W., Word, J. Q., Word, J. D., Perkins, R. A., McFarlin, K. M., Hester, B. W.,	2013	The acute toxicity of chemically and physically dispersed crude oil to key Arctic species under	*Environmental Toxicology and Chemistry 32,* 2284-2300.	**Included**	

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
	Word, L. S., Ray, C. M.		Arctic conditions during the open water season			
16	Geraudie, P., Nahrgang, J., Forget-Leray, J., Minier, C., Camus, L.	2014	*In vivo* effects of environmental concentrations of produced water on the reproductive function of polar cod (*Boreogadus saida*)	*Journal of Toxicology and Environmental Health Part A 77*(9-11), 557-573.	**Included**	
17	Grenvald, J. C., Nielsen, T. G., Hjorth, M.	2013	Effects of pyrene exposure and temperature on early development of two co-existing Arctic copepods	*Ecotoxicology 22,* 184-198.	**Included**	
18	Hansen, B. H., Nordtug, N., Altin, D., Booth, A., Hessen, K. M., Olsen, A. J.	2009	Gene expression of GST and CYP330A1 in lipid-rich and lipid poor female *Calanus finmarchicus* (Copepoda: Crustacea) exposed to dispersed crude oil	*Journal of Toxicology and Environmental Health Part A 72,* 131-139.	**Included**	
19	Hansen, B. H., Altin, D., Rørvik, S. F., Øverjordet, I. B., Olsen, A. J., Nordtug, T.	2011	Comparative study on acute effects of water accommodated fractions of an artificially weathered crude oil on *Calanus finmarchicus* and *Calanus glacialis* (Crustacea: Copepoda)	*Science of the Total Environment 409,* 704-709.	**Included**	

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
20	Hansen, B. H., Altin, D., Olsen, A. J., Nordtug, T.	2012	Acute toxicity of naturally and chemically dispersed oil on the filter-feeding copepod *Calanus finmarchicus*	*Ecotoxicology and Environmental Safety 86,* 38-46.	**Included**	[a]
21	Hansen B.H., Altin D., Rørvik S. Øverjordet, I., Jager T., and Nordtug T.	2013	Acute exposure of water soluble fractions of marine diesel on Arctic *Calanus glacialis* and boreal *Calanus finmarchicus*: Effects on survival and biomarker response	*Science of the Total Environment 449,* 276-284.	**Included**	
22	Hansen, B. H., Altin, D., Bonaunet, K., Øverjordet, I. B.	2014	Acute toxicity of eight oil spill response chemicals to temperate, boreal, and Arctic species	*Journal of Toxicology and Environmental Health Part A 77,* 495-505.	**Included**	
23	Hjorth, M. and Nielsen, T. G.	2011	Oil exposure in a warmer Arctic: Potential impacts on key zooplankton species	*Marine Biology 158*(6), 1339-1347.	**Included**	
24	Hsiao, S. I. C.	1978	Effects of crude oils on the growth of Arctic marine phytoplankton	*Environmental Pollution 17,* 93-107.	**Included**	
25	Jensen, M.H., Nielsen, T.G., Dahllöf, I.,	2008	Effects of pyrene on grazing and reproduction of *Calanus finmarchicus* and *Calanus glacialis* from Disko Bay, West Greenland	*Aquatic Toxicology 87,* 99-107.	**Included**	

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
26	Jensen, L. K., Carroll, J.	2010	Experimental studies of reproduction and feeding for two Arctic-dwelling *Calanus* species exposed to crude oil	*Aquatic Biology 10,* 261-271.	**Included**	
27	Laurel, B. J., Copeman, L. A., Iseri, P., Spencer, M. L., Hutchinson, G., Nordtug, T., Donald, C. E., Meier, S., Allan, S. E., Boyd, D. T., Ylitalo, G. M., Cameron, J. R., French, B. L., Linbo, T. L., Scholz, N. L., Incardona, J. P.	2019	Embryonic crude oil exposure impairs growth and lipid allocation in a keystone Arctic forage fish	*iScience 19,* 1101-1113.	**Included**	a
28	Lemcke, S., Holding, J., Moller, E. F., Thyrring, J., Gustavson, K. Juul-Pedersen, T., Sejr, M. K.	2018	Acute oil exposure reduces physiological process rates in Arctic phyto- and zooplankton	*Ecotoxicology 28,* 26-36.	**Included**	
29	Levin, M., E. Gebhard, L. Jasperse, J.-P. Desforges, R. Dietz, C. Sonne, I. Eulaers, A. Covaci,	2016	Immunomodulatory effects of exposure to polychlorinated biphenyls and perfluoroalkyl acids in East Greenland ringed seals (*Pusa hispida*)	*Environmental Research 151,* 244-250.	**Included**	

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
	R. Bossi and S. De Guise					
30	Moore, Dana	2016	Toxicity of salts, metals, and nitrogenous contaminants to cold-water fish under northern conditions (Chapter 2)	Unpublished doctoral dissertation	**Included**	
31	Moore, Dana	2016	Toxicity of salts, metals, and nitrogenous contaminants to cold-water fish under northern conditions (Chapter 3)	Unpublished doctoral dissertation	**Included**	
32	Moore, Dana	2016	Toxicity of salts, metals, and nitrogenous contaminants to cold-water fish under northern conditions (Chapter 4)	Unpublished doctoral dissertation	**Included**	
33	Moore, Dana	2016	Toxicity of salts, metals, and nitrogenous contaminants to cold-water fish under northern conditions (Chapter 5)	Unpublished doctoral dissertation	**Included**	
34	Nahrgang, J., Dubourg, P., Frantzen, M.	2016	Early life stages of an Arctic keystone species (*Boreogadus saida*) show high sensitivity to a water-soluble fraction of crude oil	*Environmental Pollution 218,* 605-614.	**Included**	

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
35	Nørregaard, R. D., Nielsen, T. G., Moller, E. F., Strand, J., Espersen, L., Mohl, M.	2014	Evaluating pyrene toxicity on Arctic key copepod species *Calanus hyperboreus*	*Ecotoxicology 23,* 163-174.	**Included**	
36	O'Brien, J. W.	1978	Toxicity of Prudhoe Bay crude oil to Alaskan Arctic zooplankton	*Arctic 31*(3), 219-228.	**Included**	
37	Olsen, G. H., Smit, M. G. D., Carroll, J., Jaeger, I., Smith, T., Camus, L.	2011	Arctic versus temperate comparison of risk assessment metrics for 2-methyl-naphthalene	*Marine Environmental Research 72,* 179-187.	**Included**	
38	Olsen, A. J., Nordtug, T., Altin, D., Lervik, M., Hansen, B. J.	2013	Effects of dispersed oil on reproduction in the cold water copepod *Calanus finmarchicus* (Gunnerus)	*Environmental Toxicology and Chemistry 32*(9), 2045-2055.	**Included**	[a]
39	Øverjordet, I. B., Altin, D., Berg, T., Jenssen, B. M., Gabrielsen, G. W., Hansen, B. H.	2013	Acute and sub-lethal response to mercury in Arctic and boreal calanoid copepods	*Aquatic Toxicology 155,* 160-165.	**Included**	
40	Palace, V. P., Allen-Gil, S. M., Brown, S. B., Evans, R. E., Metner, D. A., Landers, D. H., Lawrence, C. R., Klaverkamp, J. F.,	2001	Vitamin and thyroid status in arctic grayling (*Thymallus arcticus*) exposed to doses of 3,3',4,4'-tetrachlorobiphenyl that induce the phase I enzyme system	*Chemosphere 45*(2), 185-193.	**Included**	

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
	Baron, C. L., Lockhart, W. L.					
41	Pančić, M., Köhler, E., Paulsen, M. L., Toxværd, K., Lacroix, C., Le Floch, S., Hjorth, M., Nielsen, T. G.	2019	Effects of oil spill response technologies on marine microorganisms in the high Arctic	*Marine Environmental Research 151*, 104785.	**Included**	
42	Percy, J. A. and Mullin, T. C.	1975	Effects of crude oils on Arctic marine invertebrates	Beaufort Sea Technical Report No. 11. Victoria, BC: Department of the Environment.	**Included**	
43	Percy, J. A. and Mullin, T. C.	1977	Effects of crude oil on the locomotory activity of Arctic marine invertebrates	*Marine Pollution Bulletin 8*(2), 35-40.	**Included**	[a]
44	Petersen, D. G., Reichenberg, F., and Dahllöff, I.	2008	Phototoxicity of pyrene affects benthic algae and bacteria from the Arctic	*Environmental Science and Technology 42*, 1371-1376.	**Included**	
45	Riebell, P. N. and Percy, J. A.	1989	Acute toxicity of petroleum hydrocarbons to the Arctic littoral mysid, *Mysis oculata* (fabricus)	Proceedings of the Arctic Marine Oil Spill Program Technical Seminar, Calgary, AB, June 7-9,	**Included**	[a]

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
				1989. Ottawa, Ontario: Environment Canada.		
46	Rodríguez-Torres, R., Almeda, R., Kristiansen, M., Rist, S., Winding, M. S., Nielsen, T. G.	2020	Ingestion and impact of microplastics on Arctic Calanus copepods	*Aquatic Toxicology 228,* 105631.	**Included**	
47	Szczybelski, A. S., van den Heuvel-Grave, M. J., Koelmans, A. A., van den Brink, N. W.	2019	Biomarker responses and biotransformation capacity in Arctic and temperate benthic species exposed to polycyclic aromatic hydrocarbons	*Science of the Total Environment 662,* 631-638.	**Included**	
48	Thyrring, J., Juhl, B. K., Holmstrup, M.	2015	Does acute lead (Pb) contamination influence membrane fatty acid composition and freeze tolerance in intertidal blue mussels in Arctic Greenland?	*Ecotoxicology 24,* 2036-2042.	**Included**	
49	Toxværd, K., Dinh, K. V., Henriksen, O., Hjorth, M., Nielsen, T. G.	2018	Impact of pyrene exposure during overwintering of the Arctic copepod *Calanus glacialis*	*Environmental Science and Technology 52,* 10328-10336.	**Included**	[a]
50	Toxværd, K., Pančić, M., Eide, H. O., Soreide, J. E.,	2018	Effects of oil spill response technologies on the physiological	*Aquatic Toxicology 199,* 65-76.	**Included**	

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
	Lacroix, C., Le Floch, S., Hjorth, M., Nielsen, T. G.		performance of the Arctic copepod *Calanus glacialis*			
51	Toxværd, K., Dinh, K. V., Henriksen, O., Hjorth, M., Nielsen, T. G.	2019	Delayed effects of pyrene exposure during overwintering on the Arctic copepod *Calanus hyperboreus*	*Aquatic Toxicology 217,* 105332.	**Included**	[a]
52	Bejarano, A. C., Gardiner, W. W., Barron, M. G., Word, J. Q.	2017	Relative sensitivity of Arctic species to physically and chemically dispersed oil determined from three hydrocarbon measures of aquatic toxicity	*Marine Pollution Bulletin 122*, 316-322.	Excluded	Not original tox data
53	Krumhansl, K. A., Krkosek, W. H., Greenwood, M., Ragush, C., Schmidt, J., Grant, J., Barrell, J., Lu, L., Lam, B., Gagnon, G. A., Jamieson, R. C.	2015	Assessment of Arctic community wastewater impacts on marine benthic invertebrates	*Environmental Science and Technology 49*, 760-766	Excluded	Monitoring data
54	Chapman, P. M., Riddle, M. J.	2003	Missing and needed: Polar marine ecotoxicology	*Marine Pollution Bulletin 46*, 927-928	Excluded	Editorial; no original lab data
55	Chapman, P. M., Riddle, M. J.	2005	Toxic effects of contaminants in polar marine environments	*Environmental Science and*	Excluded	Editorial; no original lab data

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
				Technology 39, 200A-206A		
56	De Hoop, L., Schipper, A. M., Leuven, R. S. E. W., Huijbregts, M. A. J., Olsen, G. H., Smit, M. G. D., Hendriks, A. J.	2011	Sensitivity of polar and temperate marine organisms to oil components	*Environmental Science and Technology 45*, 9017-9023	Excluded	SSDs; no original data
57	Dietz, R., Letcher, J. R., Desforges, J.-P. et al.	2019	Current state of knowledge on biological effects from contaminants on arctic wildlife and fish	*Science of the Total Environment 696*, 133792	Excluded	Review; no original data
58	Blevin, P., Angelier, F., Tartu, S., Ruault, S., Bustamante, P., Moe, B., Bech, C., Gabrielsen, G. W., Bustnes, J. O., Chastel, O.	2016	Exposure to oxychlordane is associated with shorter teleomeres in arctic breeding kittiwakes	*Science of the Total Environment 563,* 125-130	Excluded	Original Monitoring data
59	Blevin, P., Tartu, S., Ellis, H. I. Chastel, O., Bustamante, P., Parenteau, C., Herzke, D., Angelier, F., Gabrielsen, G. W.	2017	Contaminants and energy expenditure in an Arctic seabird: organochlorine pesticides and perfluoroalkyl substances are associated with metabolic rate in a contrasted manner	*Environmental Research 157*, 118-126	Excluded	Monitoring data

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
60	Dietz, R., Sonne C. et al.	2013	What are the toxicological effects of mercury in Arctic biota?	*Science of the Total Environment 443*, 775-790	Excluded	Review; no original data; data included is Monitoring data
61	Erikstad, K. E., Moum, T., Bustnes, J. O., Reiertsen, T. K.	2010	High levels of organochlorines may affect hatching sex ratio and hatchling body mass in Arctic glaucous gulls	*Functional Ecology 25*, 289-296	Excluded	Monitoring data
62	Fisk, A. T., de Wit, C. A.	2005	An assessment of the toxicological significance of anthropogenic contaminants in Canadian Arctic wildlife	*Science of the Total Environment 351*, 57-93	Excluded	Monitoring data
63	Goutte, A., Barbraud, C. et al.	2015	Survival rate and breeding outputs in a high Arctic seabird exposed to legacy persistent organic pollutants and mercury	*Environmental Pollution 200*, 1-9	Excluded	Monitoring data
64	Hansen, B., Jager, T., Altin, D., Øverjordet, I., Olsen, A., Salaberria, I. et al.	2016	Acute toxicity of dispersed crude oil on the cold-water copepod *Calanus finmarchicus*: Elusive implications of lipid content	*Journal of Toxicology and Environmental Health Part A 79*, 13-15, 549-557	Excluded	No original tox data (further investigations of toxic mechanisms from previous studies)
65	Gustavson, L., Ciesielski, T. M., Bytinksvik, J., Styrishave, B.,	2015	Hydroxylated polychlorinated biphenyls decrease circulating	*Environmental Research 138*, 191-201	Excluded	Monitoring data

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
	Hansen, M., Lie, E., Aars, J., Jenssen, B. M.		steroids in female polar bears (*Ursus maritimus*)			
66	Gutleb, A.C., P. Cenijn, M. van Velzen, E. Lie, E. Ropstad, J.U. Skaare, T. Malmber, Å. Bergman, G.W. Gabrielsen and J. Legler	2010	*In vitro* assay shows that PCB metabolites completely saturate thyroid hormone transport capacity in blood of wild polar bears (*Ursus maritimus*).	*Environmental Science and Technology 44,* 3149-3154	Excluded	Monitoring data
67	Haarr,A., K. Hylland, N. Eckbo, G.W. Gabrielsen, J.O. Bustnes, D. Herzke, P. Blevin, O. Chastel, S.A. Hanssen, B. Moe, K. Sagerup and K. Borgå	2018	DNA damage in breeding Arctic seabirds: Baseline, sensitivity to oxidative stress and association to organohalogen contaminants	*Environmental Toxicology and Chemistry 37*, 1084-1091	Excluded	Monitoring data
68	Hallanger, I.G., E.H. Jorgensen, E. Fuglei, O. Ahlstrom, D.C.G. Muir and B.M. Jenssen	2012	Dietary contaminant exposure affects plasma testosterone, but not thyroid hormones, vitamin A, and vitamin E, in male juvenile foxes (*Vulpes lagopus*)	*Journal of Toxicology and Environmental Health A 75,* 1298-1313	Excluded	Not marine/aquatic

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
69	Hargreaves, A.L., D.P. Whiteside and G, Gilchrist	2010	Concentrations of 17 elements, including mercury, and their relationship to fitness measures in arctic shorebirds and vitamin their eggs	*Science of the Total Environment 408*, 3153-3161	Excluded	Monitoring data
70	Hegseth, Marit Nøst; Camus, Lionel; Helgason, Lisa Bjørnsdatter; Bocchetti, Raffaella; Gabrielsen, Geir Wing; Regoli, Francesco	2011a	Hepatic antioxidant responses related to levels of PCBs and metals in chicks of three Arctic seabird species	*Comparative Biochemistry and Physiology - Part C 154,* 28-35	Excluded	Monitoring data
71	Hegseth, Marit Nøst; Camus, Lionel; Gorbi, Stefania; Regoli, Francesco; Gabrielsen, Geir W	2011c	Effects of exposure to halogenated organic compounds combined with dietary restrictions on the antioxidant defense system in herring gull chicks	*Science of the Total Environment 409,* 2717-2724	Excluded	Natural exposure/not deliberate tox data/not quantifiable for tox endpoints
72	Hegseth, Marit Nøst; Gorbi, Stefania; Bocchetti, Raffaella; Camus, Lionel; Gabrielsen, Geir Wing; Regoli, Francesco.	2014	Effects of contaminant exposure and food restriction on hepatic autophagic lysosomal parameters in herring gull (*Larus argentatus*) chicks	*Comparative Biochemistry and Physiology - Part C 164*, 43-50	Excluded	Natural exposure/not deliberate tox data/not quantifiable for tox endpoints

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
73	Heinz, G., D.J. Hoffman, J. Klimstra, K. Stebbins, S. Kondrad and C. Erwin	2009	Species differences in the sensitivity of avian embryos to methylmercury	*Archives of Environmental Contamination and Toxicology 56*, 129-138	Excluded	Not Arctic
74	Helgason, L.B., J.Verreault, B.M. Braune, K. Borgå, R. Primicerio, B.M. Jenssen and G.W. Gabrielsen	2010	Relationship between persistent halogenated organic contaminants and TCDD-toxic equivalents on EROD activity and retinoid and thyroid hormone status in northern fulmars	*Science of the Total Environment 408*, 6117-6123	Excluded	Monitoring data
75	Hylland, K., Aspholm, O. O., Knutsen, J. A., and Ruus, A.	2006	Biomarkers in fish from dioxin-contaminated fjords	*Biomarkers 11,* 97-117	Excluded	Monitoring data
76	Knott, K.K., P. Schenk, S. Beyerlein, D. Boyd, G.M. Ylitalo and T.M. O'Hara	2011	Blood-based biomarkers of selenium and thyroid status indicate possible adverse biological effects of mercury and polychlorinated biphenyls in Southern Beaufort Sea polar bears	*Environmental Research 111,* 1124-1136	Excluded	Monitoring data
77	Krey, A., M. Kwan and H.M. Chan,	2014	*In vivo* and *in vitro* changes in neurochemical parameters related to mercury concentrations from specific brain regions	*Environmental Toxicology and Chemistry 33*, 2463-2471	Excluded	Monitoring data

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
			of polar bears (*Ursus maritimus*)			
78	Krey, A., S.K. Ostertag and H.M. Chan	2015	Assessment of neurotoxic effects of mercury in beluga whales (*Delphinapterus leucas*), ringed seals (*Pusa hispida*), and polar bears (*Ursus maritimus*) from the Canadian Arctic	*Science of the Total Environment 15*, 237-247	Excluded	Monitoring data
79	Kuzyk, Z.A., P.V. Hodson, S.M. Solomon and K.J. Reimer	2005	Biological responses to PCB exposure in shorthorn sculpin from Saglek Bay, Labrador	*Science of the Total Environment 351- 352*, 285-300	Excluded	Monitoring data
80	Letcher, R.J., J.O. Bustnes, R. Dietz, B.M. Jenssen, E.H. Jørgensen, C. Sonne, J.Verreault, M.M.Vijayan and G.W. Gabrielsen	2010	Exposure and effects assessment of persistent organohalogen contaminants in Arctic wildlife and fish	*Science of the Total Environment 408*, 2995- 3043	Excluded	Review on Monitoring data
81	Nomiyama, K., S. Hirakawa, A. Eguchi, C. Kanbara, D. Imaeda, J. Yoo, T. Kunisue, E.-Y. Kim, H. Iwato and S. Tanabe	2014	Toxicological assessment of polychlorinated biphenyls and their metabolites in the liver of Baikal seal (*Pusa sibirica*)	*Environmental Science and Technology 48*, 13530-13539	Excluded	Monitoring data

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
82	Bäckman, O. Pelkonen, M. Tysklind, T. Hirvi and E. Helle	2003	Contaminant exposure and effects in Baltic ringed and grey seals as assessed by biomarkers	*Marine Environmental Research 55*, 73-99	Excluded	Monitoring data
83	Oskam, I., E. Ropstad, E. Lie, A. Derocher, Ø. Wiig, E. Dahl, S. Larsen and J.U. Skaare	2004	Organochlorines affect the steroid hormone cortisol in free-ranging polar bears (*Ursus maritimus*) at Svalbard, Norway	*Journal of Toxicology and Environmental Health A 67*, 959-977	Excluded	NO ACCESS
84	Provencher, J.F., M.R. Forbes, H.L. Hennin, O.P. Love, B.M. Braune, M.L. Mallory and H.G. Gilchrist	2016	Implications of mercury and lead concentrations on breeding physiology and phenology in an Arctic bird	*Environmental Pollution 218*, 1014-1022	Excluded	Monitoring data
85	Sonne, C., P.S. Leifsson, R.Dietz, E.W. Born, R.J. Letcher, L. Hyldstrup, F.F. Riget, M. Kirkegaard, D.C.G Muir	2006b	Xenoendocrine pollutants may reduce size of sexual organs in East Greenland polar bears (*Ursus maritimus*)	*Environmental Science and Technology 40*, 5668-5674	Excluded	Monitoring data
86	Sonne, C., M. Kirkegaard, J. Jacobsen, B.M. Jenssen, R.J. Letcher and R. Dietz	2014a	Altered 25-hydroxyvitamin D3 in liver tissue from Greenland sledge dogs (*Canis familiaris*) dietary exposed to organohalogen polluted minke whale	*Ecotoxicology and Environmental Safety 104*, 403-408	Excluded	Not Arctic test species

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
			(*Balaenoptera acuterostrata*) blubber			
87	Sonne, C., R. Dietz, F.F. Rigét, R.J. Letcher, K. Munk Pedersen and B. Styrishave	2014b	Steroid hormones in blood plasma from Greenland sledge dogs (*Canis familiaris*) dietary exposed to organohalogen polluted minke whale (*Balaenoptera acutorostrata*) blubber	*Toxicology and Environmental Chemistry 96,* 273-286	Excluded	Not Arctic test species
88	Tartu, S., Lendvai, A., Blevin, P., Herzke, D., Bustamante, P., Moe, B., Gabrielsen, G.W., Bustnes, J.O. and O. Chastel	2015b	Increased adrenal responsiveness and delayed hatching date in relation to polychlorinated biphenyl (PCB exposure in Arctic breeding black-legged kittiwakes (*Rissa tridactyla*).	*General and Comparative Endocrinology 219*, 165-172	Excluded	Monitoring data
89	van den Berghe, M., L. Weijs, S. Habran, K. Das, C. Bugli and S. Pillet	2013	Effects of polychlorobiphenyls, polybromodiphenylethers, organochlorine pesticides and their metabolites on vitamin A status in lactating grey seals	*Environmental Research 120*, 18-26	Excluded	Monitoring data
90	Verboven, N., J. Verreault, R.J. Letcher, G.W.	2009	Nest temperature and incubation behavior of arctic-breeding glaucous	*Animal Behaviour 77*, 411-418	Excluded	Monitoring data

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
	Gabrielsen and N. Evans		gulls exposed to persistent organic pollutants			
91	Verreault, J., J.U. Skaare, B.M. Jenssen and G.W. Gabrielsen	2004	Effects of organochlorine contaminants on thyroid hormone levels in Arctic breeding glaucous gulls, *Larus hyperboreus*	*Environmental Health Perspectives 112*, 532-537	Excluded	Monitoring data
92	Villanger, G.D., K.M. Gabrielsen, K.M. Kovacs, C. Lydersen, E. Lie and M. Karimi,	2013	Effects of complex organohalogen contaminant mixtures on thyroid homeostasis in hooded seal (*Cystophora cristata*) mother-pup pairs.	*Chemosphere 92,* 828-842.	Excluded	Monitoring data
93	Wren, C.D., D.B. Hunter, J.F. Leatherland and P.M. Stokes	1987	The effects of polychlorinated biphenyls and methylmercury, singly and in combination, on mink. I: Uptake and toxic responses	*Archives of Environmental Contamination and Toxicology 16,* 441-447.	Excluded	Not in an Arctic context, not aquatic/marine
94	Braithwaite, L.F., Aley, M.G. and Slater, D.L.	1983	The effects of oil on the feeding mechanism of the bowhead whale.	Anchorage, AK: USDOI, MMS, Alaska OCS Region.	Excluded	Not publicly accessible
95	Carrasco-Navarro v. Jaeger I., Honlanen J., Kukkonen, J., Carroll J., and Camus L.	2015	Bioconcentration, biotransformation and elimination of pyrene in the arctic crustacean *Gammarus setosus* (Amphipoda) at two temperatures.	*Marine Environmental Research 110*, 101-109	Excluded	Monitoring data

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
96	Harvey H.R., Taylor K.A., Pie H.V., and Mitchelmore C.L.	2013	Polycyclic aromatic and aliphatic hydrocarbons in Chukchi Sea biota and sediments and their toxicological response in the Arctic cod, *Boreogadus saida*	*Deep Sea Research Part II 102,* 32-55	Excluded	Monitoring data
97	Heintz R.A., Short J.W., and Rice S.D.	1999	Sensitivity of fish embryos to weathered crude oil: Part II. Increased mortality of pink salmon (*Oncorhynchus gorbuscha*) embryos incubating downstream from weathered Exxon Valdez crude oil.	*Environmental Toxicology and Chemistry 18*(3), 494-503	Excluded	Not in an Arctic context
98	Hodson, P.V., Khan C.W., Saravanabhavan G., Clarke L., Brown R.S., Hollebone, B., Wang, Z., Short, J., Lee, K., King, T.	2007	Alkyl PAH in crude oil cause chronic toxicity to early life stages of fish	Proceedings of the Arctic and Marine Oil Spill Program, Edmonton, AB, June 5-7. Ottawa, Ontario: Environment Canada.	Excluded	Not Arctic species

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
99	Nordtug T., Olsen A., Altin D., Overrein I., Storøy W., Hansen B., and De Laender F	. 2011	Oil droplets do not affect assimilation and survival probability of first feeding larvae of northeast Arctic cod	*Science of the Total Environment 412-413*, 148-153	Excluded	Atlantic cod (*Gadus morhua*) not Arctic cod (*Boreogadus saida*); species and temperatures do not reflect exposures in an Arctic context
100	Rogstad, T. W., Sonne, C., Villanger, G. D., Ahlstom, O., Fuglei, E., Muir, D., Jorgensen, E., Munro Jenssen, B.	2017	Concentrations of vitamin A, E, thyroid and testosterone hormones in blood plasma and tissues from emaciated adult male Arctic foxes (*Vulpes lagopus*) dietary exposed to persistent organic pollutants (POPs)	*Environmental Research 154,* 284-290	Excluded	Original data but not marine/aquatic
101	Ørisland N., Englehardt F., Juck F., Hurst R., and 1 other author.	1981	Effect of crude oil on polar bears.	Environmental Studies No. 24, Northern Affairs Program, Ottawa Canada.	Excluded	Not publicly accessible
102	Smith, T. and Geraci J.	1975	The effect of contact and ingestion of crude oil on ringed seals of the Beaufort Sea	Beaufort Sea Technical Report 1975, Department of Environment: Victoria, B.C.	Excluded	Interim report without final results

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
103	Bechmann, R. K., Larsen, B. K., Taban, I. C., Hellgren, L. I., Moller, P., Sanni, S.	2010	Chronic exposure of adults and embryos of *Pandalus borealis* to oil causes PAH accumulation, initiation of biomarker responses and an increase in larval mortality	*Marine Pollution Bulletin 60*, 2087-2098	Excluded	Not in an Arctic context
104	Sundt, R. C., Pampanin, D. M., Grung, M., Barsiene, J., Ruus, A.	2011	PAH body burden and biomarker responses in mussels (*Mytilus edulis*) exposed to produced water from a North Sea oil field: Laboratory and field assessments	*Marine Pollution Bulletin 62*, 1498-1505	Excluded	Not in an Arctic context
105	Olsen, G. H., Carroll, M. L	2007	Benthic community response to petroleum associated components in Arctic versus temperate marine sediments	*Marine Biology 151,* 2167-2176	Excluded	Monitoring data
106	Paine, M. D., Leggett, W. C., McRuer, J. K., Frank, K. T.	1992	Effects of Hibernia crude oil on capelin (*Mallotus villosus*) embryos and larvae	*Marine Environmental Research 33*(3), 159-187	Excluded	Not in an Arctic context
107	Kirkegaard, M., Sonne, C., Dietz, R., Letcher, R. J., Jensen, A. L, Hansen, S. S., Jenssen, B. M., Grandjean, P.	2011	Alterations in thyroid hormone status in Greenland sledge dogs exposed to whale blubber contaminated with organohalogen compounds	*Ecotoxicology and Environmental Safety 74*, 157-163	Excluded	Not Arctic test species

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
108	Kirkegaard, M., Sonne, C., Jakobsen, J., Jenssen, B. M., Letcher, R. J., Dietz, R.	2010	Organohalogens in a whale-blubber-supplemented diet affects hepatic retinol and renal tocopherol concentrations in Greenland sled dogs (*Canis familiaris*)	*Journal of Toxicology and Environmental Health Part A Current Issues 73*, 773-786	Excluded	Not Arctic species
109	Sonne, C., Wolkers, H., Leifsson, P. S., Iburg, T., Jenssen, B. M., Fuglei, E., Ahlstrom, O., Dietz, R., Kirkegaard, M., Muir, D. C. G., Jorgensen, E. H.	2009	Chronic dietary exposure to environmental organochlorine contaminants induces thyroid gland lesions in Arctic foxes (*Vulpes lagopus*)	*Environmental Research 109*(6), 702-711	Excluded	Histopathology; not tox data
110	Sonne, C., Wolkers, H., Leifsson, P. S., Iburg, T., Jenssen, B. M., Fuglei, E., Ahlstrom, O., Dietz, R., Kirkegaard, M., Muir, D. C. G., Jorgensen, E. H.	2009	Mineral density and biomechanical properties of bone tissue from male Arctic foxes (*Vulpes lagopus*) exposed to organochlorine contaminants and emaciation	*Comparative Biochemistry and Physiology Part C 149*(1), 97-103	Excluded	Not tox data
111	Trudel, K.	1978	Effects of crude oil and crude oil/Corexit 9527 suspensions on carbon fixation by a marine phytoplankton community	Proceedings of the Arctic and Marine Oil Spill Program, Edmonton, AB, March 15-17,	Excluded	Not Arctic species or in an Arctic context

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
				1978. Ottawa, Ontario: Environment Canada.		
112	Duval, W.	1979	The sublethal effects of hydrocarbons on the bioenergetics and productivity of selected marine fauna	Proceedings of the Arctic and Marine Oil Spill Program, Edmonton, AB, March 15-17, 1979. Ottawa, Ontario: Environment Canada.	Excluded	Not Arctic species; no hard data in abstract
113	Wells, P. G. and Harris, G. W.	1980	The acute toxicity of dispersants and chemically dispersed oil	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Not Arctic species
114	Fink, R. P. and Duval, W. S.	1980	Sublethal and lethal effects of the water-soluble fraction of Prudhoe Bay crude oil on juvenile Cotto salmon (*Oncorhynchus kisutch*)	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Not Arctic species
115	Peakall, D., Gilman, A. P.	1980	The sublethal effects of oil dispersants on seabirds	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Summary of research; hard data published elsewhere
116	Andersen, J. W., Kiesser, S. L., Riley, R. G.,	1980	Toxicity of chemically dispersed oil to shrimp exposed to constant and	Proceedings of the Arctic and	Excluded	Not in an Arctic context

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
	Thomas, B. L., Bean, R. M.		decreasing concentrations in a flowing system	Marine Oil Spill Program		
117	Fink, R. P., Harwood, L. A., Duval, W. S.	1981	The sublethal effects of dispersed crude oil on an estuarine isopod	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Not Arctic species or in an Arctic context
118	Lambert, G. and Peakall, D. B.	1981	Thermoregulatory metabolism in mallard ducks exposed to crude oil and dispersant	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Not Arctic species or in an Arctic context
119	Sheppard, E. P. and Georghiou, P.	1981	The mutagenicity of Prudhoe Bay crude oil and its burn residues	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Not toxicity tests
120	Lehtinen, K.-J., Suomalainen, S., Lehtinen, C., Mattsson, J., Reiland, S., Linden, O.	1982	Physiological effects on fish chronically exposed to low levels of petroleum hydrocarbons	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Not in an Arctic context
121	Nes, H. and Norland, S.	1983	Effectiveness and toxicity of oil dispersants	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Not in an Arctic context or with Arctic species
122	Bardot, C., Bocard, C., Castaing, G., Gatellier, C.	1984	The importance of a dilution process to evaluate effectiveness and toxicity of chemical dispersants	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Not in an Arctic context or with Arctic species

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
123	Bobra, A. M., Abernethy, S., Wells, P. G. and Mackay, D.	1984	Recent toxicity studies at the University of Toronto	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Not in an Arctic context or with Arctic species
124	Lockhart, W. L., Billeck, B. N., Danell, R. W., deMarch, B. G. E., Duncan, D. A.	1984	Comparative toxicity of several oil/dispersant mixtures to representative freshwater organisms	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Not in an Arctic context or with Arctic species
125	Sveum, P. and Sendstad, E.	1985	Oil polluted seaweeds in the Arctic: Short term effects on decomposers and fate of the oil	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Not toxicity data
126	McAuliffe, C. D.	1986	Organism exposure to volatile hydrocarbons from untreated and chemically dispersed crude oils in field and laboratory	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Review of studies
127	Bakke, T.	1986	Experimental long term oil pollution in a Boreal rocky shore environment	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Not in an Arctic context
128	Klokk, T.	1986	Effects of oil and clean-up methods on shoreline vegetation	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Not in an Arctic context
129	Sveum, P.	1987	Fate and effects of dispersed and non dispersed oil on Arctic mud flats	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Monitoring data; not tox data

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
130	Crowell, M. J. and Lane, P. A.	1988	Recovery of a Nova Scotian saltmarsh during two growing seasons following experimental spills of crude oil and the dispersant Corexit 9527	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Not Arctic
131	Carter, J. and Ernst, R.	1989	Tainting in sea scallops (*Placopecten magellanicus*) exposed to the water-soluble fraction of crude oil and natural gas condensate	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Not in an Arctic context
132	Jenssen, B. M., Ekker, M., Vongraven, D., Silverstone, M.	1991	Body weight development and thermoregulation of oil-contaminated grey seal pups (*Halichoerus grypud*) at the Froan Archipelago, Norway	Proceedings of the Arctic and Marine Oil Spill Program, Vancouver, BC, June 10-11, 1991. Ottawa, Ontario: Environment Canada.	Excluded	Field study
133	Lockhart, W. L. and Metner, D. A.	1991	Oil-sensitive biomarker studies of fish from Arctic Canada	Proceedings of the Arctic and Marine Oil Spill Program, Vancouver, BC, June 10-11, 1991. Ottawa, Ontario: Environment Canada.	Excluded	Summary of studies

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
134	Lockhart, W. L. and Danell, R. W.	1992	Field and experimental tainting of Arctic freshwater fish by crude and refined petroleum products	Proceedings of the Arctic and Marine Oil Spill Program, Edmonton, AB, June 10-12, 1992. Ottawa, Ontario: Environment Canada.	Excluded	Not toxicology data
135	Ackman, R. G. and Heras, H.	1992	Tainting by short-term exposure of Atlantic salmon to water soluble petroleum hydrocarbons	Proceedings of the Arctic and Marine Oil Spill Program, Edmonton, AB, June 10-12, 1992. Ottawa, Ontario: Environment Canada.	Excluded	Not toxicology data
136	Daykin, M., Sergy, G., Aurand, D. Shigenaka, G., Wang, Z., Tang, A.	1994	Aquatic toxicity resulting from in situ burning of oil-on-water	Proceedings of the Arctic and Marine Oil Spill Program, Vancouver, BC, June 8-10, 1994. Ottawa, Ontario: Environment Canada.	Excluded	Not Arctic species or in an Arctic context
137	Vandermeulen, J. H., Vignier, V., Mossman, D.	1994	Toxicology of Hibernia crude oil in parr and smolt of Atlantic salmon (*Salmo salar*)	Proceedings of the Arctic and Marine Oil Spill Program,	Excluded	Not in an Arctic context

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
				Vancouver, BC, June 8-10, 1994. Ottawa, Ontario: Environment Canada.		
138	Coelho, G. M., Bragin, G. E., Aurand, D. V., Clark, J. R., Wright, D. A.	1995	Field and laboratory investigation of the toxicity of physically and chemically dispersed oil	Proceedings of the Arctic and Marine Oil Spill Program, Vancouver, BC, June 8-10, 1994. Ottawa, Ontario: Environment Canada.	Excluded	Not Arctic species or in an Arctic context
139	Singer, M. M., George, S., Jacobsopn, S., Weetman, L. L., Blondina, G., Tjeerdema, R. S., Aurand, D., Sowby, M. L.	1996	Evaluation of the aquatic effects of crude oil, dispersants, and their mixtures	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Overview of papers in the field
140	Lockhart, W. L., Billeck, B. N., Evans, R. E., Danell, R. A., Carneiro, J.	1996	Toxicology studies with a high-boiling lubricating oil (Esstic 46 CF) used in a hydroelectric generating station	Proceedings of the Arctic and Marine Oil Spill Program	Excluded	Not Arctic species or in an Arctic context
141	Lockart, W. L., Duncan, D. A., Billeck, B. N.,	1997	Chronic toxicity of the ‘water-soluble fraction’ of Norman Wells crude oil to juvenile fish	Proceedings of the Arctic and Marine Oil Spill Program,	Excluded	Not Arctic species or in an Arctic context

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
	Danell, R. A., Ryan, M. J.			Vancouver, BC, June 11-13, 1997. Ottawa, Ontario: Environment Canada.		
142	Blenkinsopp, S. and Sergy, G.	1997	Evaluation of the toxicity of the weathered crude oil used at the Newfoundland offshore burn experiment (NOBE) and the resultant burn residue	Proceedings of the Arctic and Marine Oil Spill Program, Vancouver, BC, June 11-13, 1997. Ottawa, Ontario: Environment Canada.	Excluded	Not Arctic species or in an Arctic context
143	Fuller, C., Bonner, J., McDonald, T., Page, C., Bragin, G., Clark, J., Aurand, D., Hernandez, A., Ernest, A.	1999	Comparative toxicity of simulated beach sediments impacted with both whole and chemical dispersions of weathered Arabian medium crude oil	Proceedings of the Arctic and Marine Oil Spill Program, Calgary, AB, June 2-4, 1997. Ottawa, Ontario: Environment Canada.	Excluded	Not Arctic species or in an Arctic context
144	Brakstad, O. G., Faksness, L.-G., Stokland, O., Altin, D., and Singsaas, I.	2000	Disappearance and biological effects of crude oils after sedimentation on subtidal soft-bottom seabed sediments: Experiments in a laboratory seabed mesocosm	Proceedings of the Arctic and Marine Oil Spill Program, Vancouver, BC, June 14-16, 2000. Ottawa, Ontario:	Excluded	Not tox data

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
				Environment Canada.		
145	Armsworthy, S., Cranford, P. J., Tremblay, G. H., and Lee, K.	2000	Chronic toxicity of Orimulsion to the sea scallop *Placopecten magellanicus*: influences on survival, feeding, digestion and growth	Proceedings of the Arctic and Marine Oil Spill Program, Vancouver, BC, June 14-16, 2000. Ottawa, Ontario: Environment Canada.	Excluded	Not in an Arctic context
146	McFarlin, K. M., Perkins, R. A., Leigh, M. B.	2012	The effects of crude oil and Corexit 9500 on the Indigenous Arctic microbial community	Proceedings of the Arctic and Marine Oil Spill Program, Vancouver, BC, June 5-7, 2012. Ottawa, Ontario: Environment Canada.	Excluded	No results (results pending)
147	McFarlin, K. M., Perkins, R. A., Leigh, M. B.	2014	Oil biodegradation by Arctic marine microorganisms	Proceedings of the Arctic and Marine Oil Spill Program, Canmore, AB, June 3-5, 2014. Ottawa, Ontario: Environment Canada.	Excluded	Not tox data

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
148	Dussauze et al.	2014	Effect of dispersed oil on fish cardiac tissue respiration: A comparison between a temperate (*Dicentrarchus labrax*) and an arctic (*Boreogadus saida*) species	Proceedings of the Arctic and Marine Oil Spill Program, Canmore, AB, June 3-5, 2014. Ottawa, Ontario: Environment Canada.	Excluded	Same study as paper assessed already (conference abstract)
149	Percy, J. A.	1977	Responses of Arctic marine benthic crustaceans to sediments contaminated with crude oil	*Environmental Pollution 13*, 1-10	Excluded	Same study as paper assessed already (conference abstract)
150	Rice, S. D., Short, J. W., Karinen, J. F.	1976	Toxicity of Cook Inlet crude oil and No. 2 fuel oil to several Alaskan marine fishes and invertebrates	In Sources, Effects, and Sinks of Hydrocarbons in the Aquatic Environment Proc. Symp., American University, Washington D. C., pp. 394-406	Excluded	Not publicly accessible
151	Rice, S. D., Short, J. W., Brodersen, C. C., Mecklenburg, T. A., Moles, D. A., Misch, C. J.,	1976	Acute toxicity and uptake-depuration studies with Cook Inlet crude oil, Prudhoe Bay crude oil, No. 2 fuel oil and several subarctic marine organisms	NWFC Processed Report, National Marine Fisheries Service. Seattle. 90 p.	Excluded	Not publicly accessible

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
	Cheatham, D. L., and Karinen, J. F.					
152	Schneider, D. E.	1980	Physiological responses of arctic epibenthic invertebrates to winter stresses and exposure to Prudhoe Bay crude oil dispersions	In Environmental Assessment of the Alaskan Continental Shelf, Annual Reports of Principal Investigators, Vol. 1, Receptors – Birds, Plankton, Littoral, Benthos. Boulder, CO: USDOC, NOAA, OCSEAP, pp. 413-474	Excluded	Not publicly accessible
153	Mageau, C., Engelhardt, F. R., Gilfillan, E. S. and Boehm, P. D.	1987	Effects of short-term exposure to dispersed oil in Arctic invertebrates	*Arctic 40(1),* 62-171	Excluded	Already assessed this study in the form of a conference paper
154	McFarlin, K. M. and Perkins, R. A.	2011	Toxicity of physically and chemically dispersed oil to selected Arctic species	2011 International Oil Spill Conference	Excluded	Already assessed this study in a paper published by a different author

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
155	Cross, W. E.	1982	In situ studies of the effects of oil and dispersed oil on primary productivity of ice algae and on under-ice amphipod communities	In Special Studies – 1981 study results, Baffin Island Oil Spill Working Report 81-10: 61 p.	Excluded	Monitoring data; not publicly accessible but data published in Cross and Martin 1987 (below) and NA
156	Cross, W. E. and Martin, C. M.	1983	*In situ* studies of the effects of oil and chemically treated oil on primary productivity of ice algae and on under-ice microfauna and macrofaunal communities	In: Special studies – 1982 study results, Baffin Island Oil Spill Working Report 82-7, 103 p.	Excluded	Monitoring data; not publicly accessible but data published in Cross and Martin 1987 (below) and NA
157	Cross, W. E. and Martin, C. M.	1987	Effects of oil and chemically treated oil on nearshore under-ice meiofauna studied in situ	*Arctic 40,* 258-265	Excluded	Monitoring data
158	Hsiao, S. I. C., Kittle, D. W., and Foy, M. G.	1978	Effects of crude oils and the oil dispersant Corexit on primary production of Arctic marine phytoplankton and seaweed	*Environmental Pollution 15,* 209-221	Excluded	Monitoring data
159	Federle, T. W., Vestal, J. R., Hater, G. R., Miller, M. C.	1979	Effects of Prudhoe bay crude oil on primary production and zooplankton in Arctic tundra thaw ponds	*Marine Environmental Research 2*(1), 3-18	Excluded	Field study

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
160	Miller, M. C., Hater, G. R., Vestal, J. R.	1977	Effect of Prudhoe crude oil on carbon assimilation by planktonic algae in an Arctic pond	In Environmental Chemistry and Cycling Processes, Pine Ridge, TN.	Excluded	Field study
161	Bergstein, P. E. and Vestal, J. R.	1978	Crude oil biodegradation in Arctic tundra ponds	*Arctic 31*(3), 153-411	Excluded	Field study
162	Jordan, M. J., Hobbie, J. E., Peterson, B. J.	1978	Effect of petroleum hydrocarbons on microbial populations in an Arctic lake	*Arctic 31*(3), 153-411	Excluded	Field study
163	Miller, M. C., Alexander, V., Barsdate, R. J.	1978	The effects of oil spills on phytoplankton in an Arctic lake and ponds	*Arctic 31*(3), 192-218	Excluded	Field study
164	Mozley, S. C. and Butler, M. G.	1978	Effects of crude oil on aquatic insects of tundra ponds	*Arctic 31*(3), 229-241	Excluded	Field study
165	Walker, D. A., Webber, P. J., Everett, K. R., Brown, J.	1978	Effects of crude and diesel spills on plant communities at Prudhoe Bay, Alaska, and the derivation of oil spill sensitivity maps	*Arctic 31*(3), 242-259	Excluded	Field study
166	McFarlin, K. M., Perkins, R. A.	2010	Evaluating the biodegradability and effects of dispersed oil using Arctic test species and conditions: Phase I activities	Proceedings of the Arctic and Marine Oil Spill Program, Halifax, NS, June 7-9, 2010. Ottawa, Ontario:	Excluded	Not tox data

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
				Environment Canada.		
167	McFarlin, K. M., Perkins, R. A.	2011	Evaluating the biodegradability and effects of dispersed oil using Arctic test species and conditions: Phase II activities	Proceedings of the Arctic and Marine Oil Spill Program, Banff, AB, June 4-6, 2011. Ottawa, Ontario: Environment Canada.	Excluded	Not tox data
168	Perkins, R. A.	2000	Selection of potential cold water marine species for testing of oil, dispersants, and chemically dispersed oil	Proceedings of the Arctic and Marine Oil Spill Program, Vancouver, BC, June 14-16, 2000. Ottawa, Ontario: Environment Canada.	Excluded	Not tox data
169	Gerlich, H. S., Holmstrup, M., Bjerregaard, P., Slotsbo, S.	2020	Mercury (Hg2+) interferes with physiological adaptations to freezing in the arctic earthworm *Enchytraeus albidus*	*Ecotoxicology and Environmental Safety 204,* 111005.	Excluded	Terrestrial
170	Reinardy, H. C., Pedersen, K. B., Nahrgang, J., Frantzen, M.	2019	Effects of mine tailings exposure on early life stages of Atlantic cod.	*Environmental Toxicology and Chemistry 38*(7), 1446-1454	Excluded	Not an Arctic test species

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
171	Caputo, S., Papale, M., Rizzo, C., Giannarelli, S., Conte, A., Moscheo, F., Graziano, M., Aspholm, P. E., Onor, M., De Domenico, E., Miserocchi, S., Michaud, L., Azzaro, M., Lo Guidice, A.	2019	Metal resistance in bacteria from contaminated Arctic sediment is driven by metal local inputs.	*Archives of Environmental Contamination and Toxicology 77*(2), 291-307	Excluded	Monitoring data
172	Brown, J., Whiteley, N. M., Bailey, A. M., Graham, H., Hop, H., Rastrick, S. P. S.	2020	Contrasting responses to salinity and future ocean acidification in Arctic populations of the amphipod *Gammarus setosus*.	*Marine Environmental Research 162,* 105176	Excluded	Not tox data
173	Brix, K. V., Baker, J., Morris, W., Ferry, K., Pettem, C., Elphick, J., Tear, L. M., Napier, R., Adzic, M., DeForest, D. K.	2021	Effects of maternally transferred egg selenium on embryo-larval survival, growth, and development in Arctic grayling (*Thymallus arcticus*).	*Environmental Toxicology and Chemistry 40*(2), 380-389	Excluded	Monitoring data
174	Rowlands, E., Galloway, T., Manno, C.	2021	A Polar outlook: Potential interactions of micro- and nano-plastic with other anthropogenic stressors.	*Science of the Total Environment 754,* 142379	Excluded	Not tox data

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
175	Camus, L., Smit, M. G. D.	2019	Environmental effects of Arctic oil spills and spill response technologies, introduction to a 5 year joint industry effort.	*Marine Environmental Research 144,* 250-254	Excluded	Not tox data
176	Fahd, F., Veitch, B., Khan, F.	2020	Risk assessment of Arctic aquatic species using ecotoxicological biomarkers and Bayesian network.	*Marine Pollution Bulletin 156,* 111212	Excluded	Modelling data
177	Pacyna-Kuchta, A. D., Jakubas, D., Frankowski, M., Polkowska, Z., Wojczulanis-Jakubas, K.	2020	Exposure of a small Arctic seabird, the little auk (*Alle alle*) breeding in Svalbard, to selected elements throughout the course of a year.	*Science of the Total Environment 732,* 139103	Excluded	Monitoring data
178	Mirimin, L., Hickey, A., Barrett, D., DeFaoite, F., Boschetti, S., Venkatesh, S., Graham, G. T.	2020	Environmental DNA detection of Arctic char (*Salvelinus alpinus*) in Irish lakes: Development and application of a species-specific molecular assay.	*Environmental DNA 2*(2), 221-233	Excluded	Monitoring data
179	Hansen, B. H., Sorensen, L., Storseth, T. R., Nepstad, R., Altin, D., Krause, D., Meier, S., Nordtug, T.	2019	Embryonic exposure to produced water can cause cardiac toxicity and deformations in Atlantic cod (*Gadus morhua*) and haddock (*Melanogrammus aeglefinus*) larvae.	*Marine Environmental Research 148*, 81-86	Excluded	Not Arctic test species

Study No.	Authors	Year	Title	Journal	Included / Excluded	Notes
180	Jain, A., Krishnan, K. P., Begum, N., Singh, A., Thomas, F. A., Gopinath, A.	2020	Response of bacterial communities from Kongsfjorden (Svalbard, Arctic Ocean) to macroalgal polysaccharide amendments.	*Marine Environmental Research 155,* 104874	Excluded	Not tox data

[a] Study includes post-exposure monitoring for recovery potential and/or latent effects observation.

Table S2 Justification of environmentally-relevant concentrations of compounds tested on Arctic species from studies assessed in this review.

Substance	Environmentally-Relevant Concentration Threshold	Justification	Study Number
OIL-RELATED CONTAMINANTS			
Chemical dispersant	< 1000 mg/L	Concentrations of dispersants in the water column after application in an oil spill are highly variable and difficult to measure. Estimates range from 5 mg/L (Negri et al., 2018) to 10 mg/L (Wells, 1984) based on measured concentrations in field experiments to 271 mg/L to 904 mg/L (Bejarano, 2018) based on worst-case exposure scenarios calculated from dispersant application rates over a depth of 0.01 m. Therefore, environmentally-relevant concentrations for this review will capture the upper end of this range.	7b, 11a, 12b, 15b, 20b, 22b, 42b
Shoreline washing agent (SWA)	Based on exposure scenario	Data pertaining to concentrations of shoreline washing agents in the water column after application are not readily available as this substance is sprayed on the shoreline. If effort has been made to represent field conditions through exposure duration or scenario, study will be scored with a 1.	22a
Crude oil (mechanically dispersed or chemically dispersed)	Based on exposure scenario	Environmentally realistic concentrations of oil in an oil spill are highly variable depending on many factors. If effort has been made to represent field conditions through exposure concentration, duration, scenario (i.e., spiked decline over constant exposure), study will be scored with a 1.	1, 3b, 3c, 3d, 3e, 5a, 7a, 8, 9, 10, 11b, 12a, 13a, 15a, 18, 19, 20a, 20c, 24, 26, 27,

Substance	Environmentally-Relevant Concentration Threshold	Justification	Study Number
	< 1000 µg/L TPH < 1000 µg/L THC	The vast majority (95%) of seawater samples taken after the Deepwater Horizon oil spill contained < 250 µg/L TPH (Wade et al., 2016); however, maximum values reached 11,400 mg/L (directly above wellhead; Sammarco et al., 2013). It has been reported that water column samples that were not directly above wellhead with concentrations above 1000 µg/L did not contain high concentrations of PAHs that correspond to these TPH measurements; thus, the TPHs measured were likely from non-petroleum, organic sources (Wade et al., 2016). For this reason, 1000 µg/L has been selected as the maximum environmentally-relevant concentration threshold for these compounds.	34, 36, 38, 41a, 41b, 41c, 42a, 43, 45a, 50
Heavy fuel oil	< 1000 µg/L TPH < 1000 µg/L THC	The vast majority (95%) of seawater samples taken after the Deepwater Horizon oil spill contained < 250 µg/L TPH (Wade et al., 2016); however, maximum values reached 11,400 mg/L (directly above wellhead; Sammarco et al., 2013). It has been reported that water column samples that were not directly above wellhead with concentrations above 1000 µg/L did not contain high concentrations of PAHs that correspond to these TPH measurements; thus, the TPHs measured were likely from non-petroleum, organic sources (Wade et al., 2016). For this reason, 1000 µg/L has been selected as the maximum environmentally-relevant concentration threshold for these compounds.	28
Pyrene	< 189 µg/L	Maximum reported concentration of PAHs measured in the water column after the Deepwater Horizon oil spill (Diercks et al., 2010).	13b, 17, 23, 25, 35, 44, 47, 49, 51
	< 1398 nM	This is the United States criteria for maximum allowable concentration of pyrene and seawater; therefore,	

Substance	Environmentally-Relevant Concentration Threshold	Justification	Study Number
		concentrations up to this value (at least) are likely to occur (Grenvald et al., 2013).	
	Study-specific for sediment (see Szczbelski et al., 2019)	Concentrations are justified within the text based on concentrations of PAHs found in sediments near where the exposure was conducted	
2-methylnaphthalene	< 189 µg/L	Maximum reported concentration of PAHs measured in the water column after the Deepwater Horizon oil spill (Diercks, A.-R. et al., 2010).	37
Naphthalene	< 189 µg/L	Maximum reported concentration of PAHs measured in the water column after the Deepwater Horizon oil spill (Diercks, A.-R. et al., 2010).	5b
Diesel (marine/Arctic)	< 1:10,000 dilution	Identified as an environmentally realistic worst-case scenario (Hansen et al., 2013).	3a, 21
Produced water	< 350 ng/L PAH	The composition of produced water is highly variable depending on many factors. Discharge is instantaneously diluted as it reaches the water the column, where modelling predicts that concentrations of PAHs (a component of the discharge) can be 25 to 350 ng/L within one kilometre from the source, and 4 to 8 ng/L within five to ten kilometres from the source (Durell et al., 2006). Environmental relevance of concentrations from studies that report measured PAH values will use these values for scoring.	4, 16
	Any dilutions of concentrate	Any dilutions of concentrate as long as concentrate was made to mimic realistic discharge.	
INORGANIC CONTAMINANTS			

Substance	Environmentally-Relevant Concentration Threshold	Justification	Study Number
Methylmercury (MeHg)	1.41 µg/g dry weight in biota	Concentrations of methylmercury naturally occurring in the Arctic environment are not this high; however, this compound bioaccumulates in the tissues of organisms feeding at high trophic levels. This value was obtained as the maximum from concentrations of **total mercury** found in eggs of arctic seabirds (black-legged kittiwakes, northern fulmars, and thick-billed murres) between 1975 and 2003 (Braune, 2007).	2
	1 µg/g wet weight in biota	This value was obtained as the maximum from concentrations of **total mercury** found in arctic fish muscle between 1990 and 2009 (AMAP, 2011). Wet weight concentrations were not reported for birds.	
Total mercury (THg)	In solution: 2.9 ng/L	This is the highest value that was measured in surface seawater as a result of snowmelt deposition in an Arctic region (Dommergue et al., 2010). The season in which this phenomenon occurs coincides with major biological events (i.e., phytoplankton blooms followed by zooplankton blooms); therefore, Arctic species will potentially encounter these concentrations of THg in seawater.	39
Inorganic Mercury	Study-specific (see Frouin et al., 2012)	Bioaccumulation of Hg in marine mammals can reach high levels and be absorbed into every component of the body (organs, blood, tissues, and cells; Das et al., 2016). Chronic, low exposures over time can result in exceedingly high concentrations in biota. Due to the difficulty of defining a single threshold value for these complex interactions, thresholds for these compounds, and others where bioaccumulation is a significant factor to consider, will be based on study-specific scenarios.	14a

Substance	Environmentally-Relevant Concentration Threshold	Justification	Study Number
Organic Mercury	Study-specific (see Frouin et al., 2012)	Bioaccumulation of Hg in marine mammals can reach high levels and be absorbed into every component of the body (organs, blood, tissues, and cells; Das et al.2016). Chronic, low exposures over time can result in exceedingly high concentrations in biota. Due to the difficulty of defining a single threshold value for these complex interactions, thresholds for these compounds, and others where bioaccumulation is a significant factor to consider, will be based on study-specific scenarios.	14b
Lead	< 3000 µg Pb/g dry weight	Range of possible uptake into soft tissue for the common mussel *Mytulis edulis* (Schulz-Baldes, 1974).	48
Microplastics	14,400 particles/L	The greatest concentration of microplastics found in Arctic snow (Bergmann et al., 2019).	46
Cl^{-} [a,b]	640 mg/L	CCME Water Quality Guideline for the Protection of Aquatic Life (CCME, 2011).	30
Aluminum [a]	21.2 µg/L	Maximum concentration measured in the field by Moore et al., 2016.	31, 32
Arsenic [a]	5.0 µg/L	CCME Water Quality Guideline for the Protection of Aquatic Life (CCME, 2001).	31, 32
Cadmium [a]	1.0 µg/L	CCME Water Quality Guideline for the Protection of Aquatic Life (CCME, 2014).	31, 32
Chromium [a]	8.9 µg/L	CCME Water Quality Guideline for the Protection of Aquatic Life (CCME, 1999a).	31, 32
Copper [a]	0.6 µg/L	Maximum concentration measured in the field by Moore et al., 2016.	31, 32
Nickel [a]	18.5 µg/L	Maximum concentration measured in the field by Moore et al., 2016.	31, 32
Zinc [a]	37 µg/L	CCME Water Quality Guideline for the Protection of Aquatic Life (CCME, 2018).	31, 32

Substance	Environmentally-Relevant Concentration Threshold	Justification	Study Number
NO_2 [a]	5.83 mg/L	Maximum concentration measured in the field by Moore et al., 2016.	33
NO_3 [a]	310 mg/L	Maximum concentration measured in the field by Moore et al., 2016.	33
NH_3 [a]	0.019 mg/L	CCME Water Quality Guideline for the Protection of Aquatic Life (CCME, 2010).	33
NH_4 [a]	200 mg/L	Maximum concentration measured in the field by Moore et al., 2016.	33
Na_2SO_4	0.350 g/L	Maximum observed WQO measured in the field by Moore et al., 2016.	30
Na_2MoO_4	73 µg/L	CCME Water Quality Guideline for the Protection of Aquatic Life (CCME, 1999b).	30
PHENOLS AND FLAME RETARDANTS			
PCBs	In biota: 600 mg/kg lipid weight	PCBs measured in blubber have reached 574 mg/kg lipid weight (Buckman et al., 2011); therefore, the environmentally realistic concentration for this exposure scenario captures the upper end of this measurement.	6a, 6b, 40, 29a, 29b, 29c, 29d
	Based on exposure scenario	Environmentally realistic concentrations of PCBs are highly variable depending on many factors. If effort has been made to represent what may be observed in nature through exposure concentration, duration, derivation of compounds, or scenario, study will be scored with a 1.	
Phenol	< 1000 µg/L	The maximum environmentally realistic concentration of total petroleum hydrocarbons, of which phenol is one when measured during an oil spill (see justification for crude oil)	45b

[a] Values are for acute exposure in freshwater environments to mirror the exposure scenario from this study.
[b] Value is for all compounds containing chloride; see Moore et al., 2016.

Table S3 Description of purity and/or grade of substances that have been tested on Arctic species from studies assessed in this review that do not have easily measured purities or grades.

Substance	Purity/Grade
Chemical dispersant	Artificially weathered (AWT), naturally weathered (NWT), or fresh (FR) will be reported as the purity. Weathering process and time included. Score will be given depending on reporting of dispersant type.
Oil or fuel	Weathered (WT) or fresh (FR) will be reported as the purity. Weathering process and time included. Score will be given depending on reporting of oil or fuel type.

Table S4 Relevant environmental parameters upon which scoring was based for organisms used as test species in studies assessed in this review.

Test Type	Organism Type	Minimum Parameters Identified
In vivo	Fish, crustaceans, other invertebrates, molluscs	Temperature and dissolved oxygen
	Primary producers	Temperature, light, photoperiod, pH
	Birds	Temperature
	Mammals	Temperature and dissolved oxygen
In vitro	All	Temperature, media type

Table S5 Key knowledge gaps identified in the present study and recommendations for addressing each.

Knowledge Gap	Recommendation
Lack of standard toxicity and recovery test methods for Arctic species.	Strive toward developing standard laboratory tests (whole organism and *in vitro*), and in the absence of standard tests provide a thorough description of what was modified from standard methods.
Lack of toxicity data for Arctic-focused risk assessments in general, especially for some ecologically important groups.	Generate toxicity data using organisms that are likely the most sensitive first (i.e., algae). Emphasize the development and refinement of models and surrogates (e.g., temperate species) using these data and confirm protection with ongoing toxicity testing.
Lack of diversity in compounds tested on Arctic organisms.	Design repeatable and reliable experiments using standard reference toxicants so that results can be compared across space and time.
Inconsistencies in oil spill related research has led to a lack of comparability between studies, Arctic or otherwise.	Focus research on developing, validating, and refining laboratory methods and predictive effects models.
Lack of highly relevant sublethal and chronic endpoints assessed in Arctic toxicity testing.	Prioritize the incorporation of these endpoints into future Arctic toxicity test protocols.

Table S6 Chemical composition of each media type used in the optimization of culture conditions phase of the present study.

	Molar Concentration in Final Medium (M)	
Compound	Guillard's f/2 + Si	Harrison's medium + Si
$NaNO_3$	8.82×10^{-4}	5.49×10^{-4}
$NaH_2PO_4 \cdot H_2O$	3.62×10^{-5}	2.18×10^{-5}
$Na_2SiO_3 \cdot 9H_20$	1.06×10^{-4}	1.06×10^{-4}
$FeCl3 \cdot 6H20$	1.17×10^{-5}	6.56×10^{-6}
$Na_2EDTA \cdot 2H_2O$	1.17×10^{-5}	1.49×10^{-5}
$CuSO_4 \cdot 5H_2O$	3.93×10^{-8}	
$Na_2MoO_4 \cdot 2H_2O$	2.60×10^{-8}	5.20×10^{-7}
$ZnSO_4 \cdot 7H_2O$	7.65×10^{-8}	2.54×10^{-7}
$CoCl_2 \cdot 6H_2O$	4.20×10^{-8}	
$MnCl_2 \cdot 4H_2O$	9.10×10^{-7}	
Thiamine HCl (Vitamin B_1)	2.96×10^{-7}	2.97×10^{-4}
Biotin (Vitamin H)	2.05×10^{-9}	4.09×10^{-6}
Cyanocobalamin (Vitamin B_{12})	3.69×10^{-10}	1.47×10^{-6}
$MnSO_4 \cdot 4H_2O$		2.42×10^{-6}
$CoSO_4 \cdot 7H_2O$		5.69×10^{-8}
H_3BO_3		6.15×10^{-5}
Na_2SeO_3		1.00×10^{-9}

Table S7 Measured water quality parameters for each trial conducted in the present study. Test substances are Cu = copper, Zn = zinc, and MN = 1-methylnaphthalene.

Trial ID	Species	Time Point	Treatment	Temp. (°C)	Salinity (psu)	DO (% sat)	pH
			Test Solutions				
1	*N. frigida*	Pre-Exposure	Control	2.6	31.9	93.7	8.607
			Cu - Low	2.2	32.4	94.8	8.693
			Cu - Low Medium	1.9	32.2	95.3	8.646
			Cu - Medium	1.8	32.4	95.1	8.598
			Cu - Medium High	1.7	32.2	95.4	8.563
			Cu - High	1.7	32.4	96.5	8.528
			Zn - Low	2.0	32.0	96.8	8.622
			Zn - Low Medium	1.9	32.1	96.2	8.638
			Zn - Medium	1.7	32.0	95.7	8.642
			Zn - Medium High	1.7	32.1	95.8	8.631
			Zn - High	1.7	32.1	95.7	8.618
			pH Units				
1	*N. frigida*	Pre-Exposure	Control	NM	NM	NM	8.737
			Cu - Low	NM	NM	NM	8.764
			Cu - Low Medium	NM	NM	NM	8.660
			Cu - Medium	NM	NM	NM	8.631
			Cu - Medium High	NM	NM	NM	8.566
			Cu - High	NM	NM	NM	8.604
			MN - Low	NM	NM	NM	8.602
			MN - Low Medium	NM	NM	NM	8.566
			MN - Medium	NM	NM	NM	8.559
			MN - Medium High	NM	NM	NM	8.543
			MN - High	NM	NM	NM	8.524
			Zn - Low	NM	NM	NM	8.740
			Zn - Low Medium	NM	NM	NM	8.736
			Zn - Medium	NM	NM	NM	8.790
			Zn - Medium High	NM	NM	NM	8.809
			Zn - High	NM	NM	NM	8.794
		Post-Exposure	Control	NM	NM	NM	8.388
			Cu - Low	NM	NM	NM	8.443
			Cu - Low Medium	NM	NM	NM	8.463
			Cu - Medium	NM	NM	NM	8.399
			Cu - Medium High	NM	NM	NM	8.182
			Cu - High	NM	NM	NM	8.097
			MN - Low	NM	NM	NM	8.561
			MN - Low Medium	NM	NM	NM	8.372
			MN - Medium	NM	NM	NM	8.057
			MN - Medium High	NM	NM	NM	8.116
			MN - High	NM	NM	NM	7.967
			Zn - Low	NM	NM	NM	8.425
			Zn - Low Medium	NM	NM	NM	8.451
			Zn - Medium	NM	NM	NM	8.406
			Zn - Medium High	NM	NM	NM	8.317
			Zn - High	NM	NM	NM	8.151
			Test Solutions				
2	*N. frigida*		Control	2.4	30.9	99.8	8.502

Trial ID	Species	Time Point	Treatment	Temp. (°C)	Salinity (psu)	DO (% sat)	pH
			Cu - Low	2.4	30.7	97.7	8.503
			Cu - Low Medium	2.5	30.8	97.9	8.499
			Cu - Medium	2.6	30.8	98.3	8.480
			Cu - Medium High	2.6	30.7	98.1	8.496
			Cu - High	2.4	30.8	98.1	8.506
			Zn - Low	2.4	30.8	98.0	8.492
			Zn - Low Medium	2.4	30.8	98.4	8.504
			Zn - Medium	2.6	30.9	98.9	8.510
		Pre-Exposure	Zn - Medium High	2.4	30.6	98.7	8.506
			Zn - High	2.5	30.6	98.8	8.439
			pH Units				
			Control	NM	NM	NM	8.499
			Cu - Low	NM	NM	NM	8.498
			Cu - Low Medium	NM	NM	NM	8.501
			Cu - Medium	NM	NM	NM	8.481
			Cu - Medium High	NM	NM	NM	8.497
		Pre-Exposure	Cu - High	NM	NM	NM	8.508
			Zn - Low	NM	NM	NM	8.493
			Zn - Low Medium	NM	NM	NM	8.500
			Zn - Medium	NM	NM	NM	8.516
			Zn - Medium High	NM	NM	NM	8.505
			Zn - High	NM	NM	NM	8.431
			Control	NM	NM	NM	8.499
			Cu - Low	NM	NM	NM	8.291
2	*N. frigida*		Cu - Low Medium	NM	NM	NM	8.397
			Cu - Medium	NM	NM	NM	8.524
			Cu - Medium High	NM	NM	NM	8.209
			Cu - High	NM	NM	NM	8.145
			MN - Low	NM	NM	NM	8.288
		Post-Exposure	MN - Low Medium	NM	NM	NM	8.164
			MN - Medium	NM	NM	NM	8.108
			MN - Medium High	NM	NM	NM	7.990
			MN - High	NM	NM	NM	8.066
			Zn - Low	NM	NM	NM	8.308
			Zn - Low Medium	NM	NM	NM	8.573
			Zn - Medium	NM	NM	NM	8.425
			Zn - Medium High	NM	NM	NM	8.299
			Zn - High	NM	NM	NM	8.288
			Test Solutions				
			Control	3.2	31.6	94.0	8.516
			Cu - Low	2.3	30.8	92.1	8.624
			Cu - Low Medium	2.1	31.3	91.6	8.668
			Cu - Medium	0.6	31.1	91.7	8.673
			Cu - Medium High	1.6	31.2	91.8	8.666
3	*N. frigida*	Pre-Exposure	Cu - High	1.9	31.3	91.2	8.617
			Zn - Low	1.7	31.3	90.6	8.589
			Zn - Low Medium	1.7	31.6	91.1	8.592
			Zn - Medium	1.7	31.5	91.5	8.571
			Zn - Medium High	1.8	31.3	92.4	8.476
			Zn - High	1.9	31.2	92.1	8.454
			pH Units				
3	*N. frigida*		Control	NM	NM	NM	8.527

Trial ID	Species	Time Point	Treatment	Temp. (°C)	Salinity (psu)	DO (% sat)	pH
		Pre-Exposure	Control	NM	NM	NM	8.498
			Cu - Low	NM	NM	NM	8.532
			Cu - Low Medium	NM	NM	NM	8.579
			Cu - Medium	NM	NM	NM	8.613
			Cu - Medium High	NM	NM	NM	8.626
			Cu - High	NM	NM	NM	8.524
			MN - Low	NM	NM	NM	8.601
			MN - Low Medium	NM	NM	NM	8.475
			MN - Medium	NM	NM	NM	8.479
			MN - Medium High	NM	NM	NM	8.535
			MN - High	NM	NM	NM	8.544
			Zn - Low	NM	NM	NM	8.583
			Zn - Low Medium	NM	NM	NM	8.594
			Zn - Medium	NM	NM	NM	8.529
			Zn - Medium High	NM	NM	NM	8.489
			Zn - High	NM	NM	NM	8.494
		Post-Exposure	Control	NM	NM	NM	8.257
			Control	NM	NM	NM	8.314
			Cu - Low	NM	NM	NM	8.360
			Cu - Low Medium	NM	NM	NM	8.361
			Cu - Medium	NM	NM	NM	8.332
			Cu - Medium High	NM	NM	NM	8.210
			Cu - High	NM	NM	NM	8.178
			MN - Low	NM	NM	NM	8.357
			MN - Low Medium	NM	NM	NM	7.651
			MN - Medium	NM	NM	NM	7.604
			MN - Medium High	NM	NM	NM	7.999
			MN - High	NM	NM	NM	7.495
			Zn - Low	NM	NM	NM	8.355
			Zn - Low Medium	NM	NM	NM	8.372
			Zn - Medium	NM	NM	NM	8.369
			Zn - Medium High	NM	NM	NM	8.296
			Zn - High	NM	NM	NM	8.204
Test Solution							
1	*S. costatum*	Pre-Exposure	Control	19.6	31.7	99.8	8.349
			Cu - Low	19.7	31.7	99.6	8.327
			Cu - Low Medium	19.6	31.7	98.4	8.364
			Cu - Medium	19.6	31.7	99.1	8.399
			Cu - Medium High	19.7	31.6	97.3	8.384
			Cu - High	19.8	31.7	100.4	8.300
			Zn - Low	19.8	31.6	99.8	8.314
			Zn - Low Medium	19.8	31.6	97.1	8.360
			Zn - Medium	19.8	31.6	98.9	8.355
			Zn - Medium High	19.7	31.7	99.0	8.298
			Zn - High	19.6	31.6	101.1	8.313
pH Units							
1	*S. costatum*	Pre-Exposure	Control	NM	NM	NM	8.341
			Cu - Low	NM	NM	NM	8.319
			Cu - Low Medium	NM	NM	NM	8.361
			Cu - Medium	NM	NM	NM	8.399
			Cu - Medium High	NM	NM	NM	8.385
			Cu - High	NM	NM	NM	8.304

Trial ID	Species	Time Point	Treatment	Temp. (°C)	Salinity (psu)	DO (% sat)	pH
			Zn - Low	NM	NM	NM	8.320
			Zn - Low Medium	NM	NM	NM	8.355
			Zn - Medium	NM	NM	NM	8.361
			Zn - Medium High	NM	NM	NM	8.301
			Zn - High	NM	NM	NM	8.310
		Post-Exposure	Control	NM	NM	NM	8.836
			Control	NM	NM	NM	8.973
			Cu - Low	NM	NM	NM	8.994
			Cu - Low Medium	NM	NM	NM	8.852
			Cu - Medium	NM	NM	NM	8.320
			Cu - Medium High	NM	NM	NM	8.277
			Cu - High	NM	NM	NM	8.258
			MN - Low	NM	NM	NM	7.932
			MN - Low Medium	NM	NM	NM	8.177
			MN - Medium	NM	NM	NM	8.462
			MN - Medium High	NM	NM	NM	7.894
			MN - High	NM	NM	NM	8.003
			Zn - Low	NM	NM	NM	8.958
			Zn - Low Medium	NM	NM	NM	8.924
			Zn - Medium	NM	NM	NM	8.272
			Zn - Medium High	NM	NM	NM	8.204
			Zn - High	NM	NM	NM	8.241
Test Solutions							
2	*S. costatum*	Pre-Exposure	Control	19.8	32.9	98.6	8.633
			Cu - Low	19.1	32.8	97.1	8.587
			Cu - Low Medium	19.1	32.8	97.7	8.599
			Cu - Medium	19.0	33.0	97.5	8.574
			Cu - Medium High	19.0	32.8	97.5	8.573
			Cu - High	19.0	33.2	97.0	8.570
			Zn - Low	9.0	32.6	97.5	8.621
			Zn - Low Medium	19.0	32.7	97.3	8.610
			Zn - Medium	19.0	33.2	97.5	8.606
			Zn - Medium High	19.0	32.7	97.4	8.563
			Zn - High	19	32.6	97.2	8.532
pH Units							
2	*S. costatum*	Pre-Exposure	Control	NM	NM	NM	8.591
			Control	NM	NM	NM	8.677
			Cu - Low	NM	NM	NM	8.615
			Cu - Low Medium	NM	NM	NM	8.647
			Cu - Medium	NM	NM	NM	8.575
			Cu - Medium High	NM	NM	NM	8.600
			Cu - High	NM	NM	NM	8.600
			MN - Low	NM	NM	NM	8.575
			MN - Low Medium	NM	NM	NM	8.535
			MN - Medium	NM	NM	NM	8.532
			MN - Medium High	NM	NM	NM	8.520
			MN - High	NM	NM	NM	8.513
			Zn - Low	NM	NM	NM	8.638
			Zn - Low Medium	NM	NM	NM	8.610
			Zn - Medium	NM	NM	NM	8.592
			Zn - Medium High	NM	NM	NM	8.544
			Zn - High	NM	NM	NM	8.552

Trial ID	Species	Time Point	Treatment	Temp. (°C)	Salinity (psu)	DO (% sat)	pH
			Control	NM	NM	NM	8.693
			Control	NM	NM	NM	8.905
			Cu - Low	NM	NM	NM	8.910
			Cu - Low Medium	NM	NM	NM	8.832
			Cu - Medium	NM	NM	NM	8.831
			Cu - Medium High	NM	NM	NM	8.126
			Cu - High	NM	NM	NM	8.361
		Post-Exposure	MN - Low	NM	NM	NM	8.349
			MN - Low Medium	NM	NM	NM	8.472
			MN - Medium	NM	NM	NM	8.346
			MN - Medium High	NM	NM	NM	8.221
			MN - High	NM	NM	NM	7.962
			Zn - Low	NM	NM	NM	8.912
			Zn - Low Medium	NM	NM	NM	8.895
			Zn - Medium High	NM	NM	NM	8.077
			Zn - High	NM	NM	NM	8.169
			Test Solutions				
			Control	17.5	32.7	97.0	8.546
			Cu - Low	17.0	32.8	96.7	8.350
			Cu - Low Medium	17.1	32.9	96.8	8.431
			Cu - Medium	17.2	33.3	96.5	8.505
			Cu - Medium High	17.5	33.5	97.3	8.457
3	*S. costatum*	Pre-Exposure	Cu - High	17.3	33.3	96.8	8.432
			Zn - Low	17.2	33.6	96.6	8.317
			Zn - Low Medium	17.3	33.5	96.6	8.295
			Zn - Medium	17.4	33.4	96.8	8.282
			Zn - Medium High	17.5	32.8	96.7	8.399
			Zn - High	17.5	3.3	96.6	8.498
			pH Units				
			Control	NM	NM	NM	8.948
			Control	NM	NM	NM	8.850
			Cu - Low	NM	NM	NM	8.685
			Cu - Low Medium	NM	NM	NM	8.566
			Cu - Medium	NM	NM	NM	8.698
			Cu - Medium High	NM	NM	NM	8.179
			Cu - High	NM	NM	NM	8.351
			MN - Low	NM	NM	NM	7.814
		Pre-Exposure	MN - Low Medium	NM	NM	NM	7.830
			MN - Medium	NM	NM	NM	7.626
			MN - Medium High	NM	NM	NM	7.936
3	*S. costatum*		MN - High	NM	NM	NM	7.656
			Zn - Low	NM	NM	NM	8.642
			Zn - Low Medium	NM	NM	NM	8.636
			Zn - Medium	NM	NM	NM	8.410
			Zn - Medium High	NM	NM	NM	8.173
			Zn - High	NM	NM	NM	8.211
			Control	NM	NM	NM	8.528
			Control	NM	NM	NM	8.533
		Post-Exposure	Cu - Low	NM	NM	NM	8.393
			Cu - Low Medium	NM	NM	NM	8.475
			Cu - Medium	NM	NM	NM	8.549
			Cu - Medium High	NM	NM	NM	8.468

Trial ID	Species	Time Point	Treatment	Temp. (°C)	Salinity (psu)	DO (% sat)	pH
			Cu - High	NM	NM	NM	8.446
			MN - Low	NM	NM	NM	8.455
			MN - Low Medium	NM	NM	NM	8.373
			MN - Medium	NM	NM	NM	8.419
			MN - Medium High	NM	NM	NM	8.388
			MN - High	NM	NM	NM	8.452
			Zn - Low	NM	NM	NM	8.313
			Zn - Low Medium	NM	NM	NM	8.334
			Zn - Medium	NM	NM	NM	8.358
			Zn - Medium High	NM	NM	NM	8.478
			Zn - High	NM	NM	NM	8.522

Note: NM = not measured.

Table S8 All EC50 values calculated in the present study and associated *p*-values. Data was analyzed and significance of relationships assessed using the *drc* package in R. Statistical significance ($p < 0.05$) is indicated with an asterisk (*). Test substances are Cu = copper, Zn = zinc, MN = 1-methylnaphthalene.

Species	Test Substance	Trial Day	EC50 (yield) (µg/L)	Standard Error	Lower Limit (95% CI)	Upper Limit (95% CI)	*p*-value	
N. frigida	Cu	8	1.41	40.39	-85.23	88.04	0.97	
			1084.48	304.46	431.49	1737.48	3.10×10^{-3}	*
			NA	NA	NA	NA	9.99×10^{-1}	
		10	1035.29	173.22	663.77	1406.80	3.39×10^{-5}	*
			663.43	246.34	135.07	1191.78	0.02	*
			66.15	NA	NA	NA	NA	
		12	1833.70	1387.00	-1141.10	4808.50	0.21	
			1032.80	101.10	815.90	1249.60	7.18×10^{-8}	*
			1239.64	804.75	-486.38	2965.67	0.15	
		14	1483.75	437.37	545.69	2421.82	4.40×10^{-3}	*
			1204.26	237.54	694.78	1713.74	2.00×10^{-4}	*
			1071.09	420.89	168.38	1973.81	0.02	*
		16	1111.93	94.44	909.36	1314.49	1.19×10^{-8}	*
			1020.02	43.77	926.16	1113.89	1.34×10^{-12}	*
			1043.17	232.26	545.01	1541.32	5.00×10^{-4}	*
	MN	8	889.64	NA	NA	NA	0.33	
			2048.80	10.00	2027.30	2070.20	2.20×10^{-16}	*
			271.60	287.90	-345.89	889.09	0.36	
		10	88.50	30.31	23.50	153.50	0.01	*
			97.39	20.32	53.81	140.98	3.00×10^{-4}	*
			356.94	793.05	-1343.98	2057.86	0.66	
		12	125.95	28.46	64.92	186.98	6.00×10^{-4}	*
			177.06	20.45	133.20	220.92	5.39×10^{-7}	*
			113.69	10.72	90.70	136.68	4.49×10^{-8}	*
		14	150.31	599.72	-1135.95	1436.57	0.81	
			327.87	103.28	106.35	549.39	6.80×10^{-3}	*
			386.67	294.41	-244.76	1018.11	0.21	
		16	126.96	49.30	21.22	232.70	0.02	*
			48.47	7.21	33.00	63.94	9.79×10^{-6}	*
			424.12	104.14	200.76	647.48	1.10×10^{-3}	*
	Zn	8	1350.30	NA	NA	NA	NA	
			1066.73	883.59	-828.39	2961.85	0.25	
			991.64	937.97	-1007.23	2990.50	0.31	

Species	Test Substance	Trial Day	EC50 (yield) (µg/L)	Standard Error	Lower Limit (95% CI)	Upper Limit (95% CI)	*p*-value	
			2337.40	NA	NA	NA	NA	
		10	1604.80	2226.80	-3171.30	6380.80	0.48	
			807874	10.00	807852	807895	2.20×10^{-16}	*
			663.34	81.80	487.89	838.78	1.17×10^{-6}	*
		12	1358.86	550.23	178.73	2539.00	0.03	*
			8817.50	44579.40	-86795.70	104430.80	0.85	
			1385.74	575.57	151.26	2620.22	0.03	*
		14	1076.27	364.44	294.63	1857.91	0.01	*
			1074.52	178.41	691.87	1457.17	3.00×10^{-5}	*
			1185.73	346.97	441.56	1929.91	4.20×10^{-3}	*
		16	1355.00	1901.00	-2722.30	5432.30	0.49	
			1472.00	6519.50	-12510.90	15454.80	0.82	
S. costatum	Cu		2129.03	334.47	1411.67	2846.39	1.75×10^{-5}	*
		4	2801.01	236.97	2292.75	3309.26	1.14×10^{-8}	*
			2412.81	199.14	1985.70	2839.92	8.26×10^{-9}	*
			2156.70	140.80	1854.70	2458.70	3.86×10^{-10}	*
		5	2339.90	108.43	2107.33	2572.46	3.83×10^{-12}	*
			2478.08	141.54	2174.51	2781.66	6.48×10^{-11}	*
			2285.37	115.87	2036.86	2533.88	1.30×10^{-11}	*
		6	2258.12	180.45	1871.10	2645.15	5.45×10^{-9}	*
			4953.20	6415.70	-8807.00	18713.50	0.45	
			2193.00	33.07	2122.08	2263.93	2.20×10^{-16}	*
		7	2424.19	187.39	2022.27	2826.10	3.55×10^{-9}	*
			2634.29	36.84	2555.28	2713.31	2.20×10^{-16}	*
	MN		236.67	102.64	16.54	456.81	0.04	*
		4	292.57	196.59	-129.08	714.22	0.16	
			137.36	6.51	123.38	151.33	5.23×10^{-12}	*
			224.63	187.31	-177.11	626.37	0.25	
		5	91.56	21.76	44.88	138.24	9.00×10^{-4}	*
			38.93	17.17	2.10	75.77	0.04	*
			474.38	189.96	66.95	881.81	0.03	*
		6	268.49	43.20	175.85	361.14	2.25×10^{-5}	*
			47.91	6.93	33.05	62.76	7.12×10^{-6}	*
			299.51	166.48	-57.55	656.57	0.09	
		7	212.12	37.02	132.73	291.51	5.20×10^{-5}	*
			182.68	2.54	177.23	188.12	2.20×10^{-16}	*
	Zn	4	2255.31	95.25	2051.03	2459.60	1.08×10^{-12}	*

Species	Test Substance	Trial Day	EC50 (yield) (µg/L)	Standard Error	Lower Limit (95% CI)	Upper Limit (95% CI)	*p*-value	
			2501.77	46.02	2403.06	2600.47	2.20×10^{-16}	*
			2123.49	34.22	2050.10	2196.88	2.20×10^{-16}	*
			2511.63	216.42	2047.45	2975.81	1.44×10^{-8}	*
		5	2215.68	122.41	1953.13	2478.23	4.14×10^{-11}	*
			2386.50	51.61	2275.81	2497.19	2.20×10^{-16}	*
			2639.89	366.92	1852.92	3426.86	4.60×10^{-6}	*
		6	2221.42	146.14	1907.97	2534.86	4.27×10^{-10}	*
			2384.16	100.91	2167.72	2600.60	1.11×10^{-12}	*
			2059.04	115.73	1810.83	2307.26	5.22×10^{-11}	*
		7	2550.17	19.53	2507.19	2593.15	2.20×10^{-16}	*
			2701.19	42.21	2610.67	2791.71	2.20×10^{-16}	*

5.2 Figures

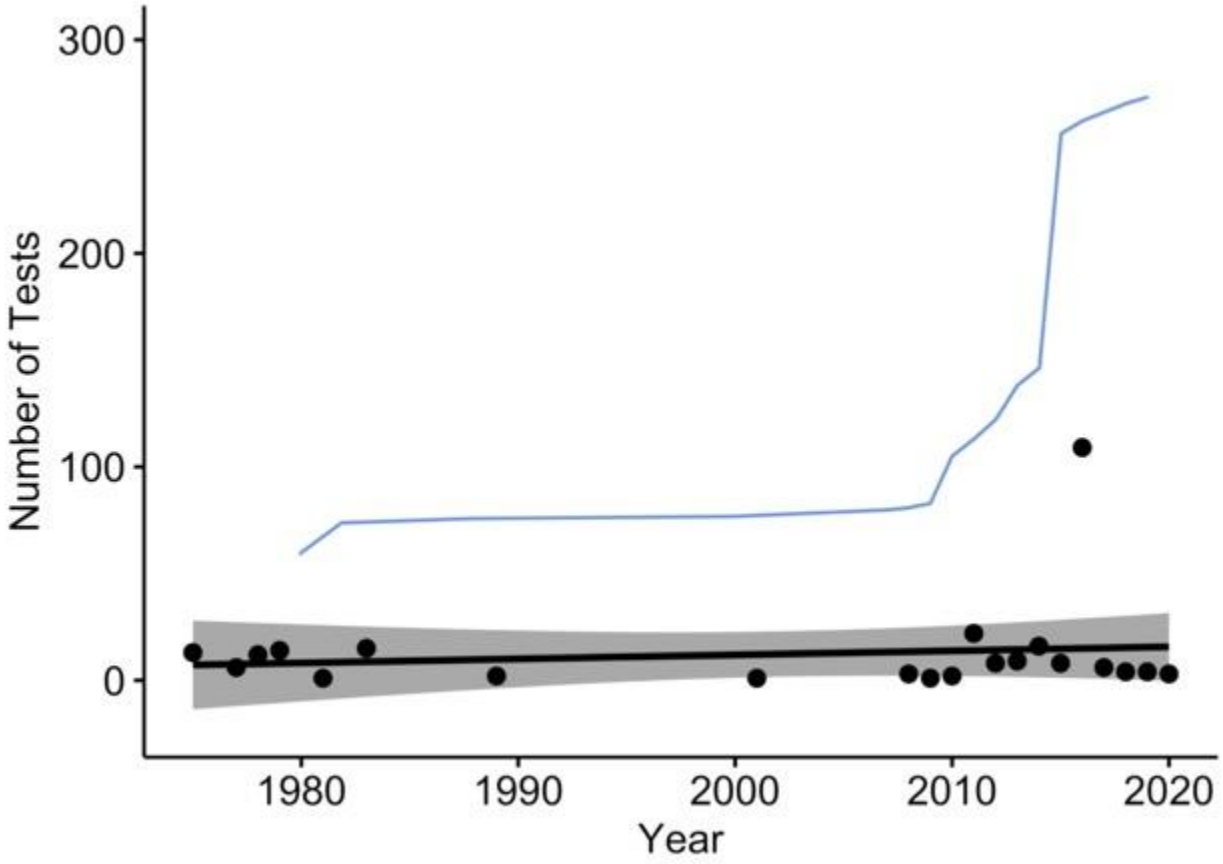

Figure S1 Number of unique toxicity tests (*n* = 253) that met inclusion criteria outlined in the present review from 1975 to 2020. Blue line is cumulative output of toxicity tests over time. Spearman's correlation coefficient (rho) = 0, *p* = 1.

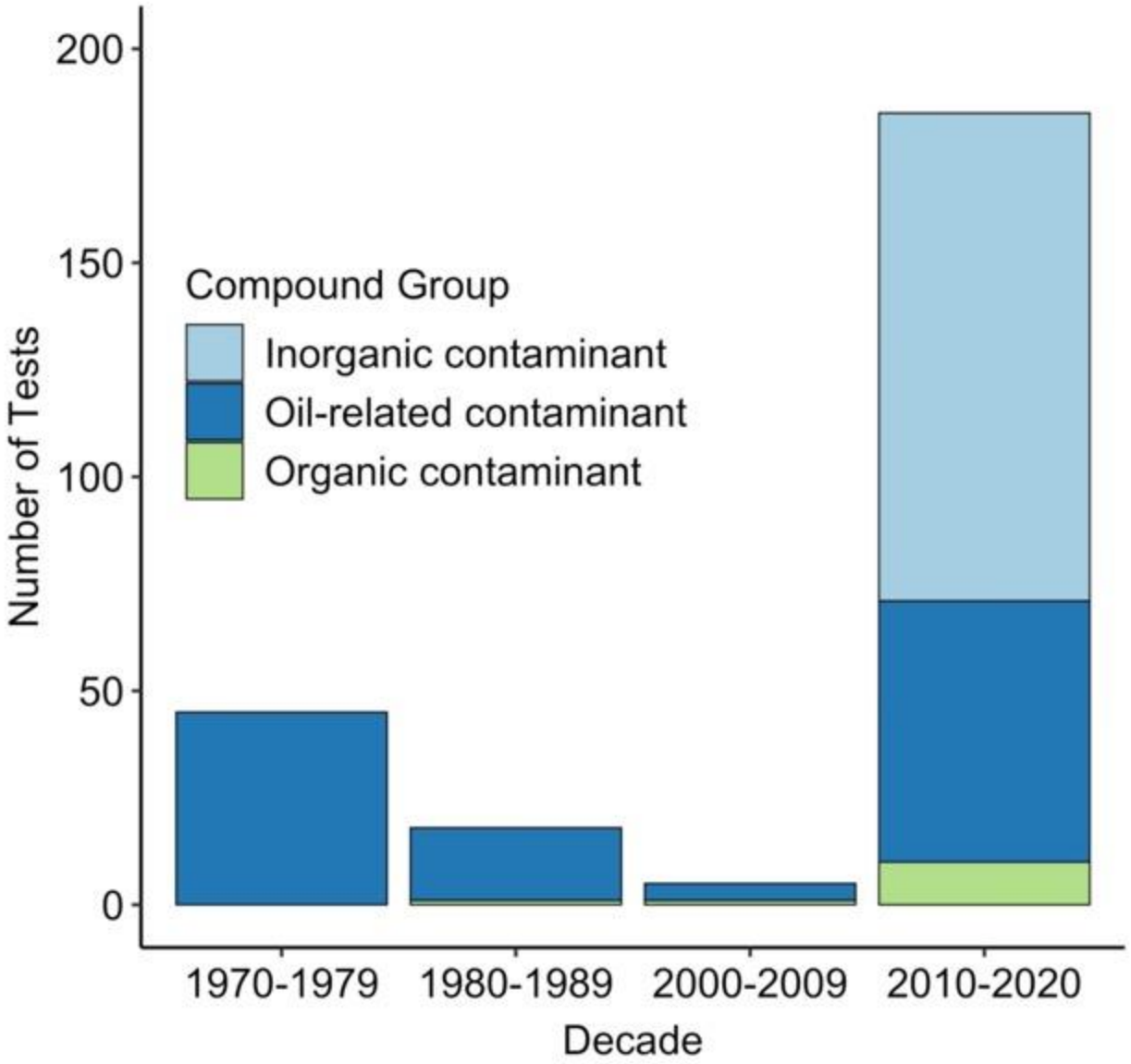

Figure S2 Proportion of each compound group observed among all unique toxicity tests captured in this review (*n* = 253 as of June 2021) by decade.

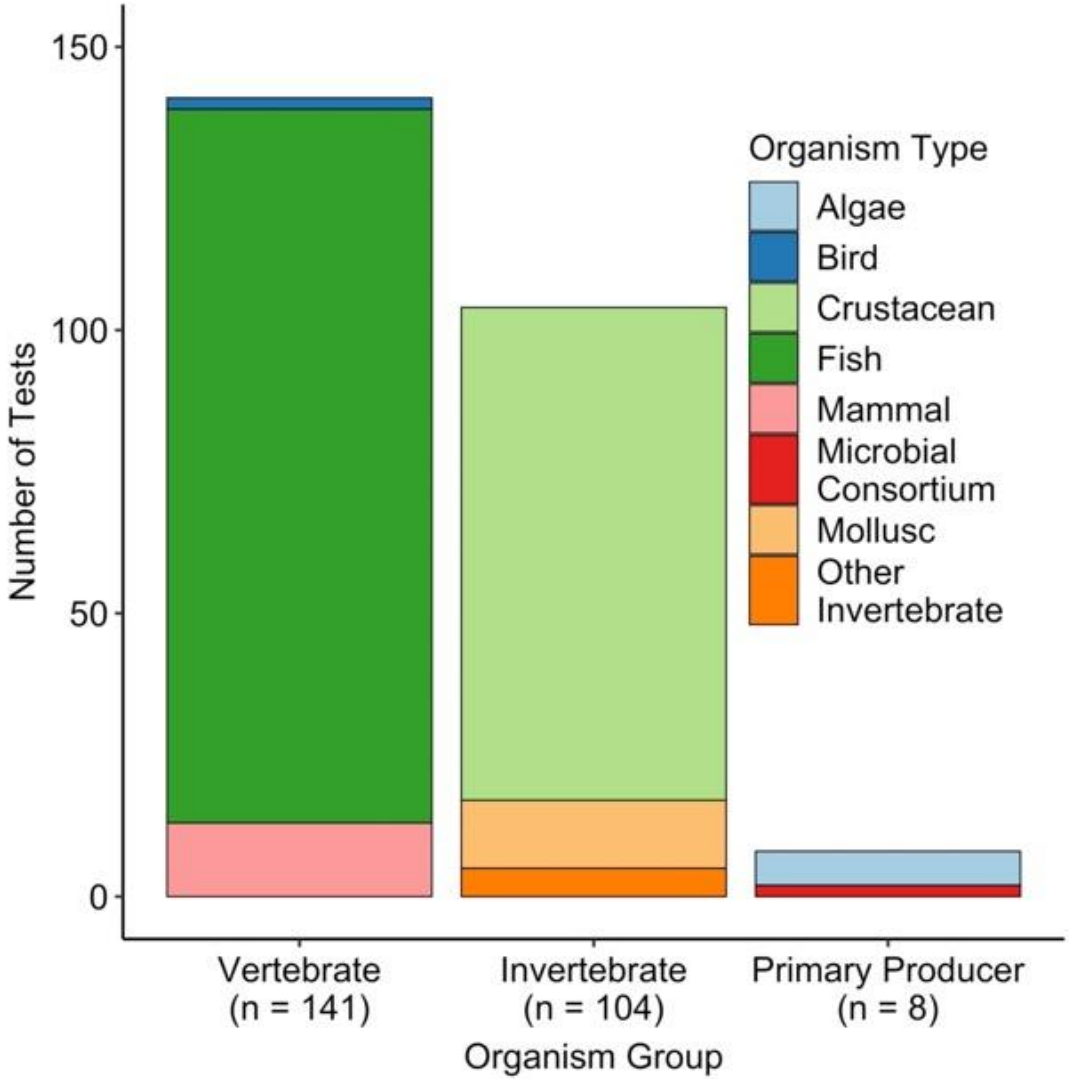

Figure S3 Number of unique Arctic species toxicity tests captured in this review (*n* = 253 as of June 2021) by organism group, with organism type within bars (number below each bar is total number of unique tests).

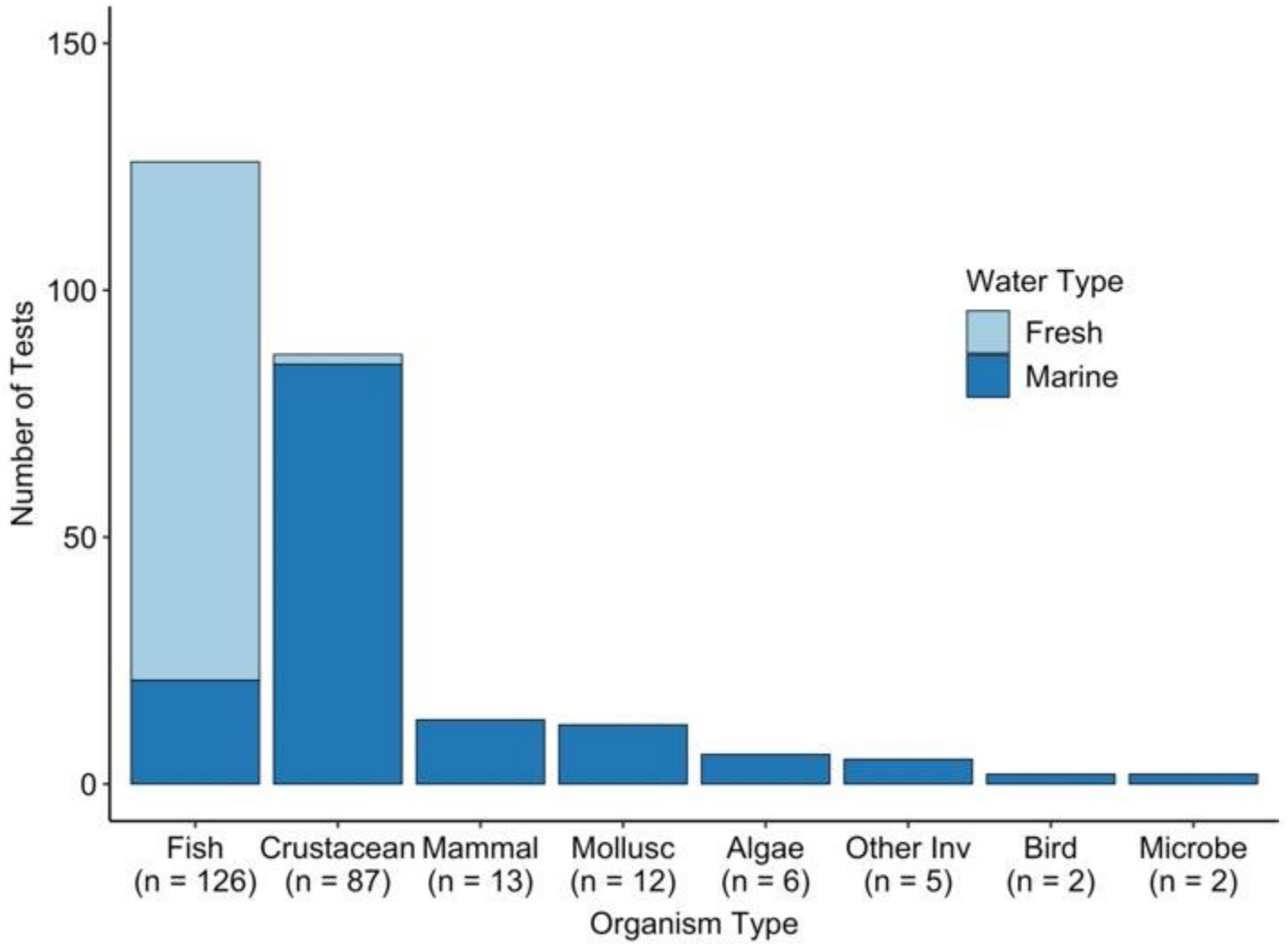

Figure S4 Proportion of marine and freshwater unique toxicity tests captured in this review (*n* = 253 as of June 2021) by organism type (number below each bar is total number of unique tests).

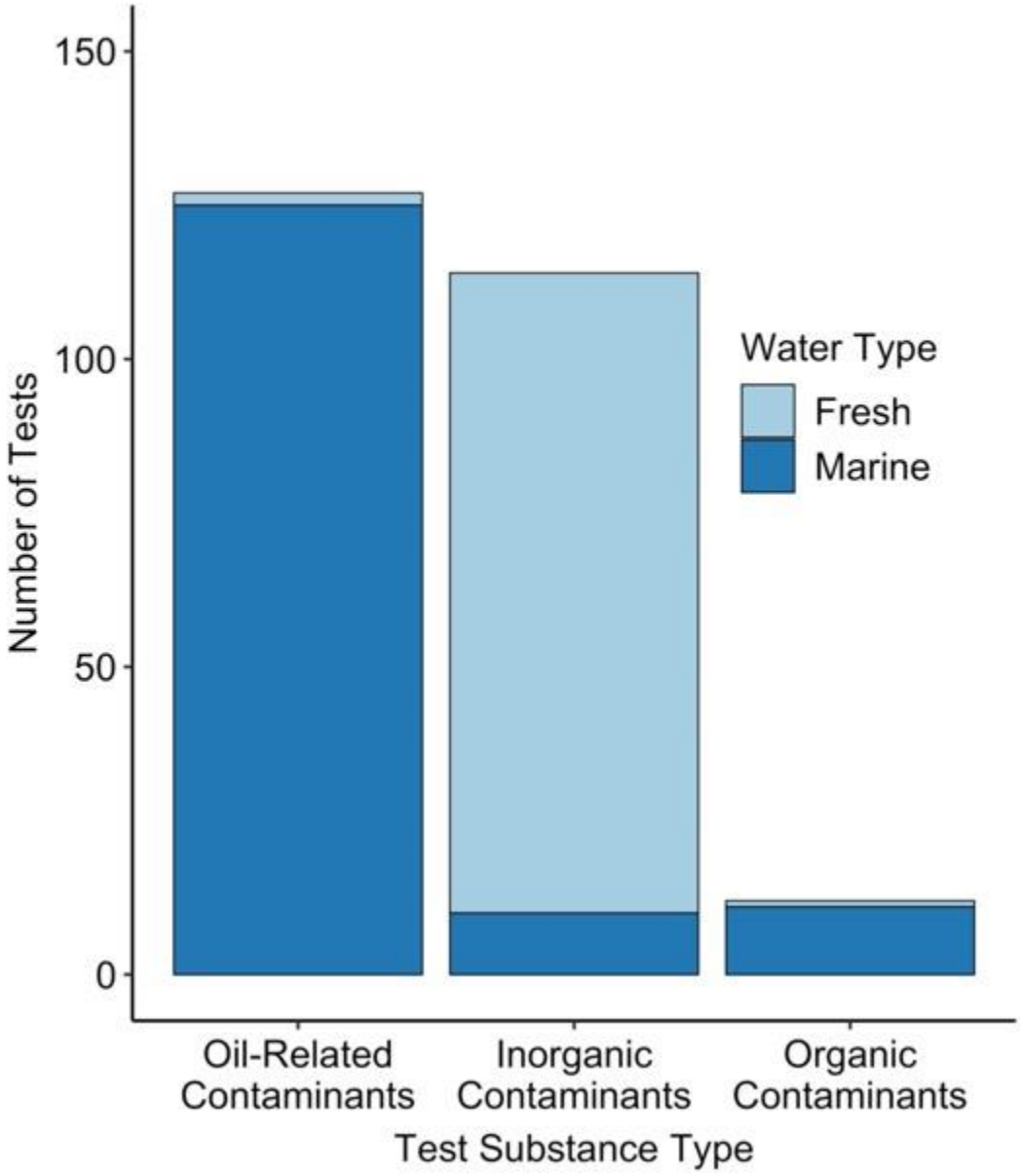

Figure S5 Proportion of marine and freshwater unique toxicity tests captured in this review (*n* = 253 as of June 2021) by test compound group.

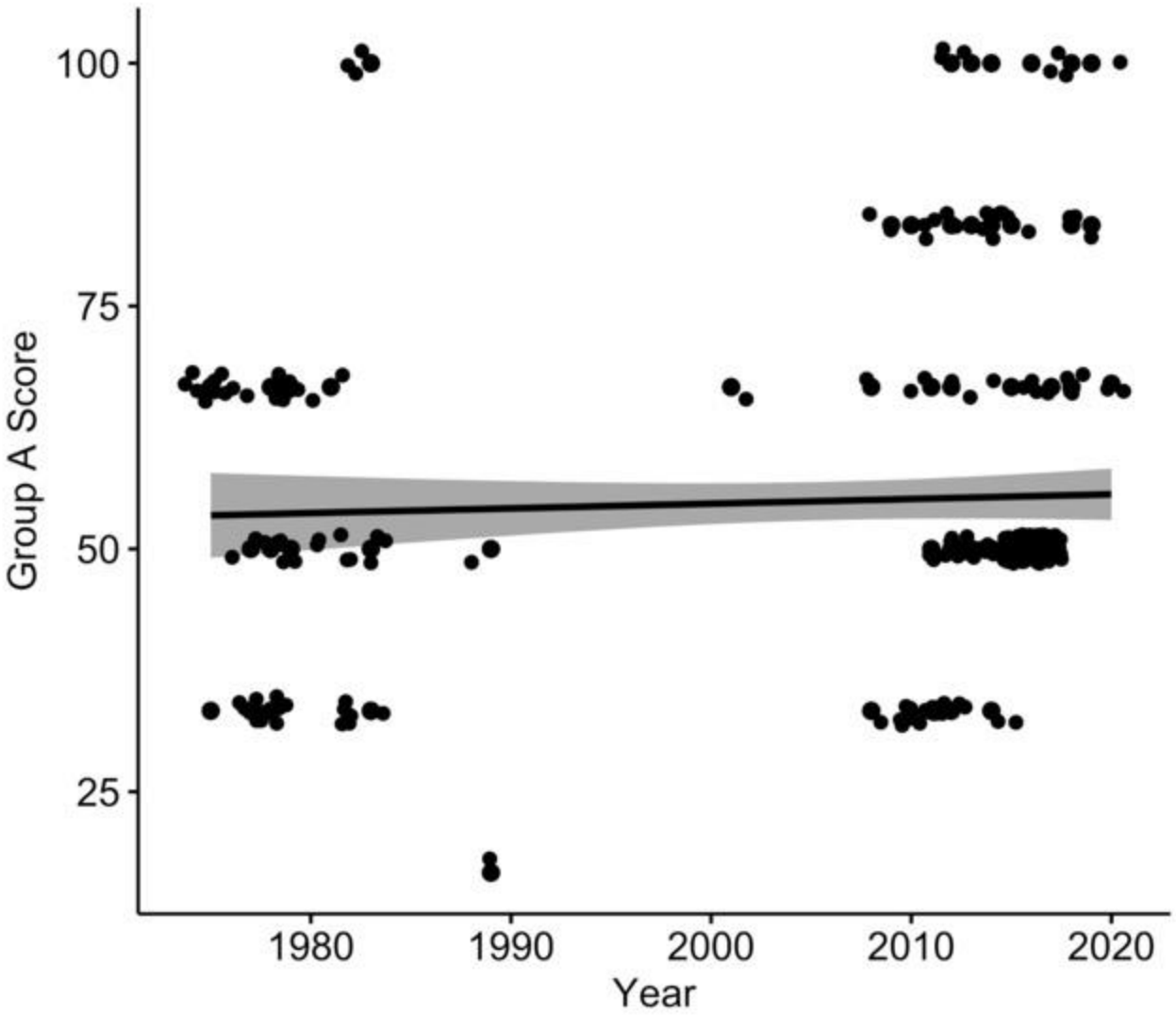

Figure S6 Scores (as percentage) over time (*n* = 253) for criterion group A (test substance) across all studies assessed in the present review. Spearman's correlation coefficient (rho) = 0.09, *p* = 0.16.

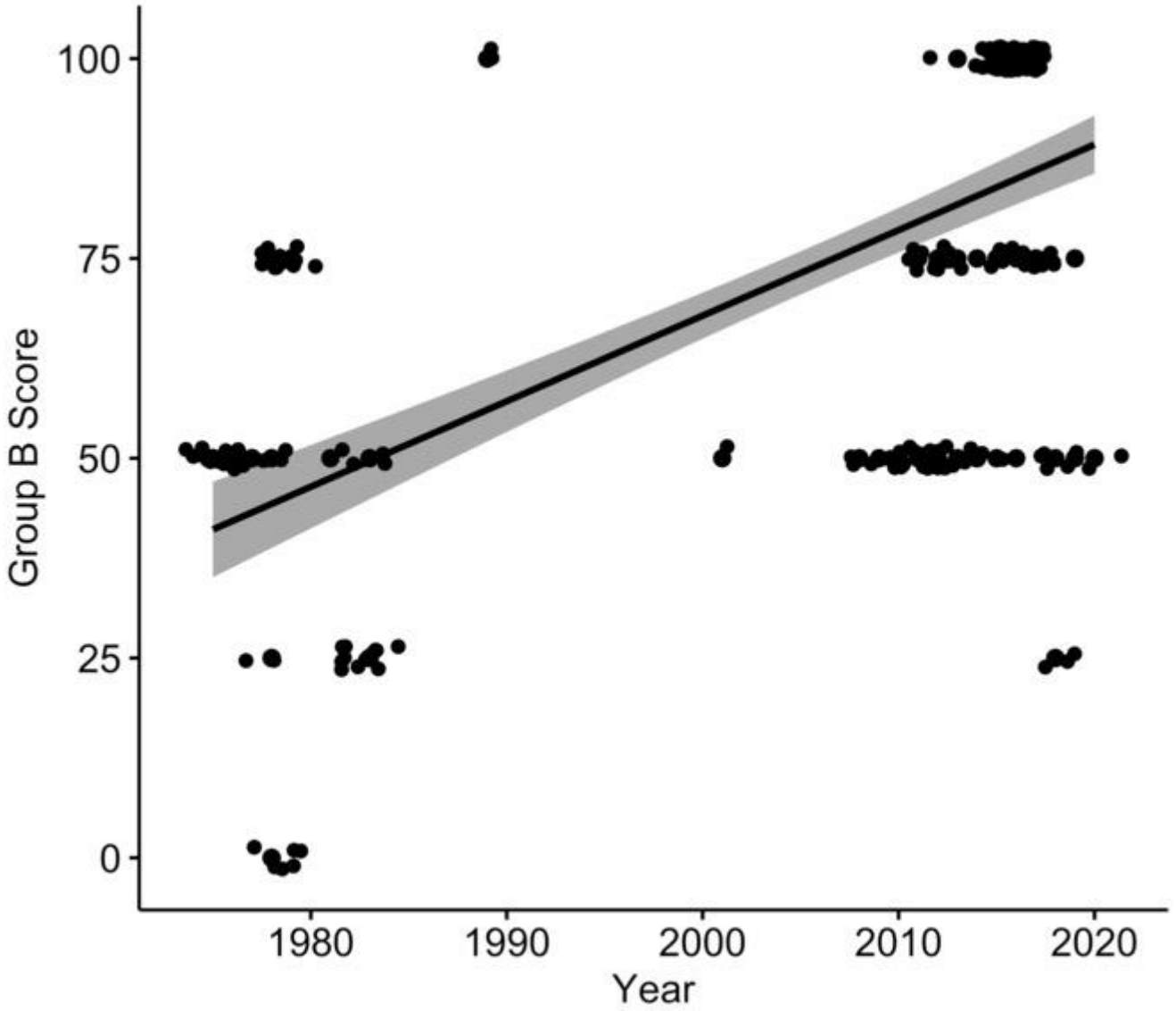

Figure S7 Scores (as a percentage) over time (n = 253) for criterion group B (test organism and test system) across all studies assessed in the present review. Spearman's correlation coefficient (rho) = 0.65, $p < 0.001$.

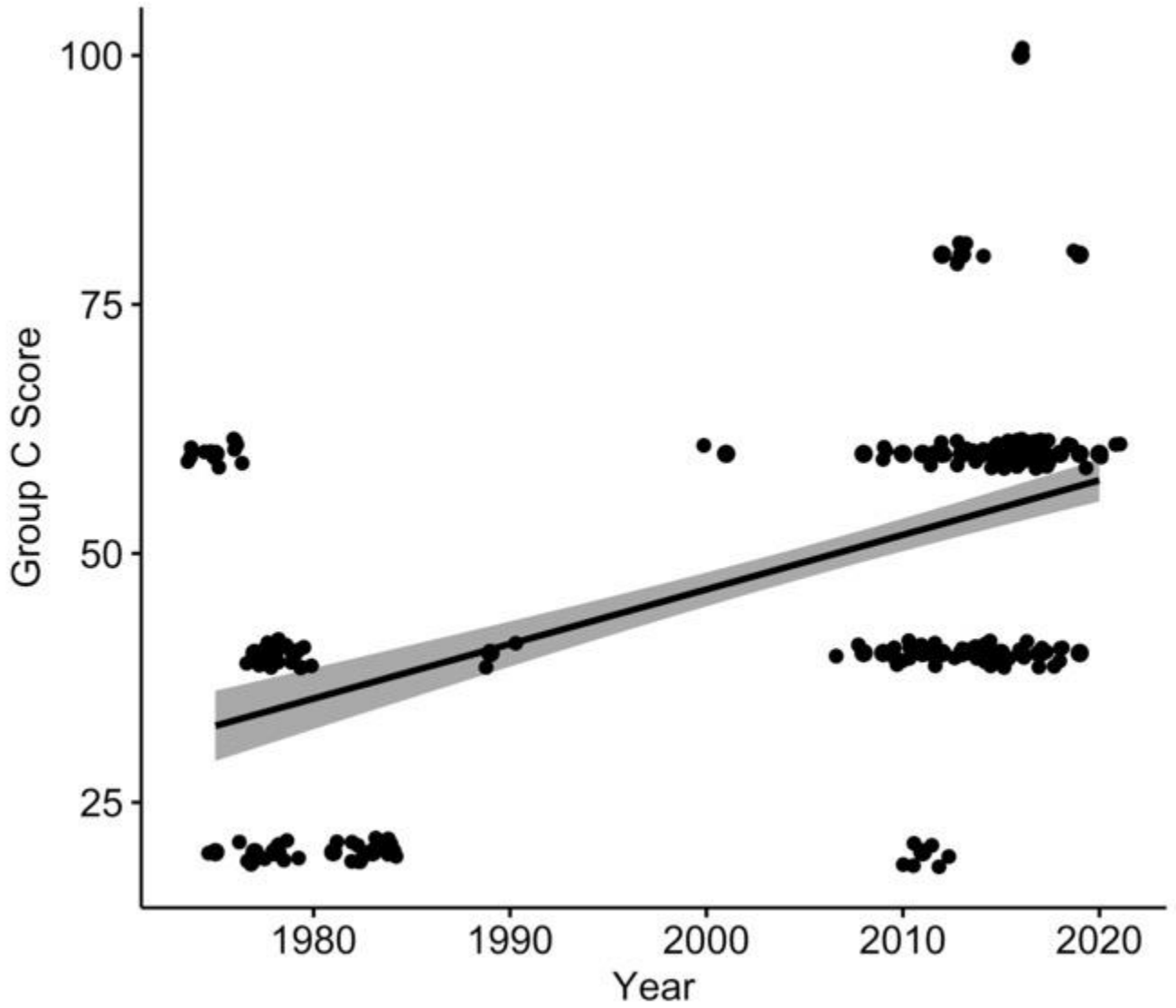

Figure S8 Scores (as a percentage) over time (n = 253) for criterion group C (test design, statistics, and results) across all studies assessed in the present review. Spearman's correlation coefficient (rho) = 0.56, $p < 0.001$.

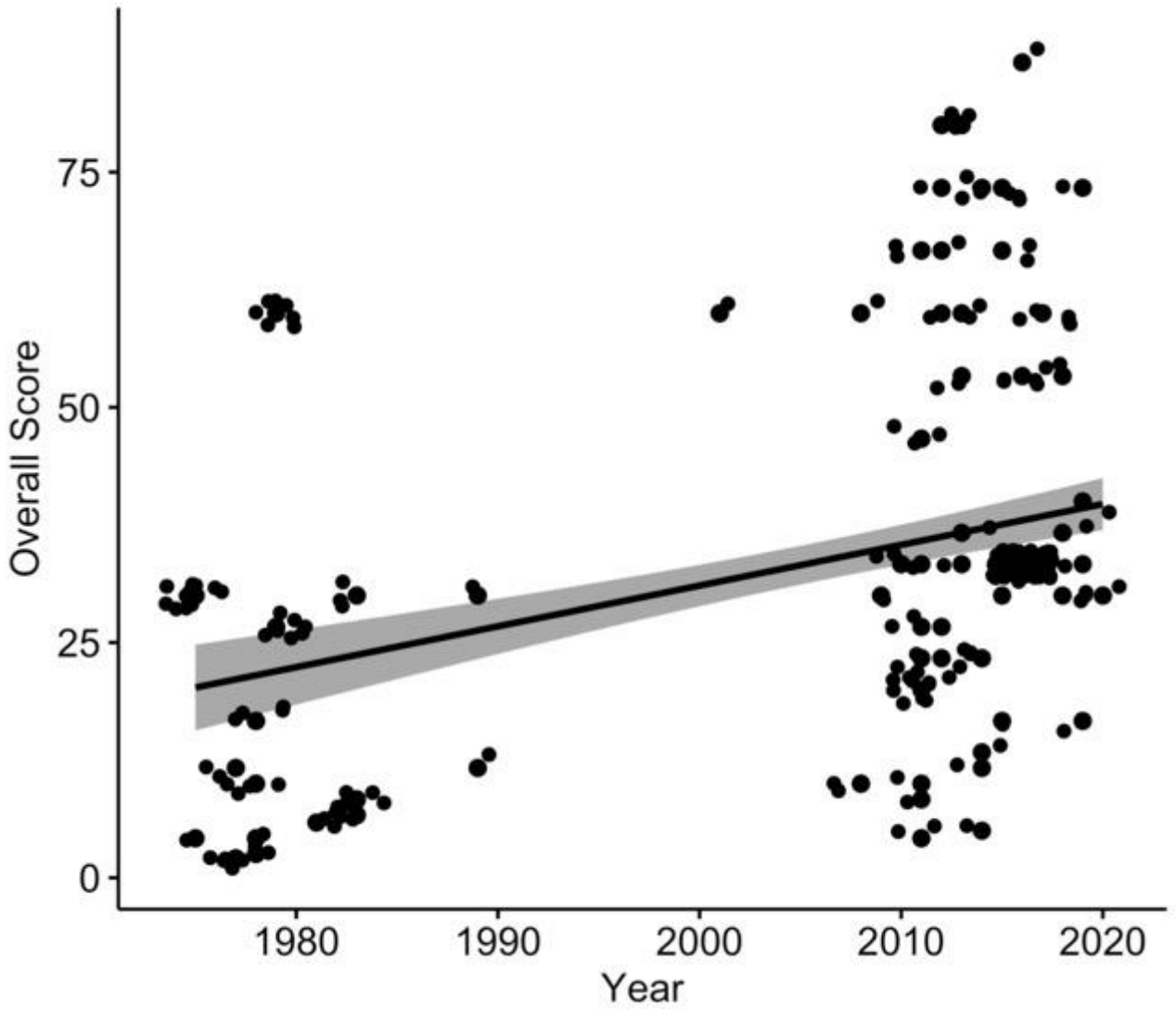

Figure S9 Overall scores (as a percentage) over time (*n* = 253) across all studies assessed in the present review. Spearman's correlation coefficient (rho) = 0.5, $p < 0.001$.

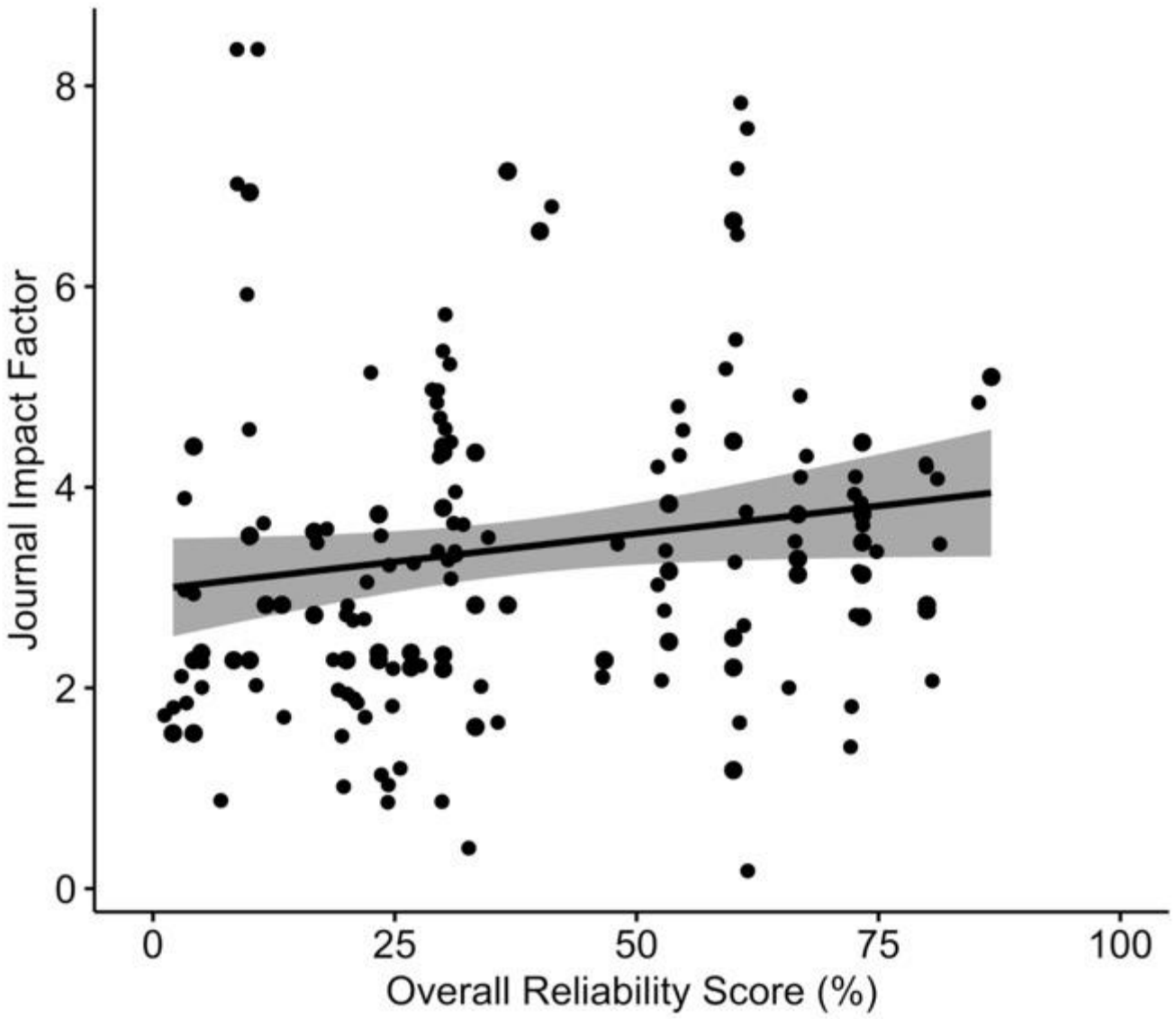

Figure S10 Relationship between overall score and the journal impact factor at the time of publication across all studies assessed in the present review. Spearman's correlation coefficient (rho) = 0.33, $p < 0.001$.

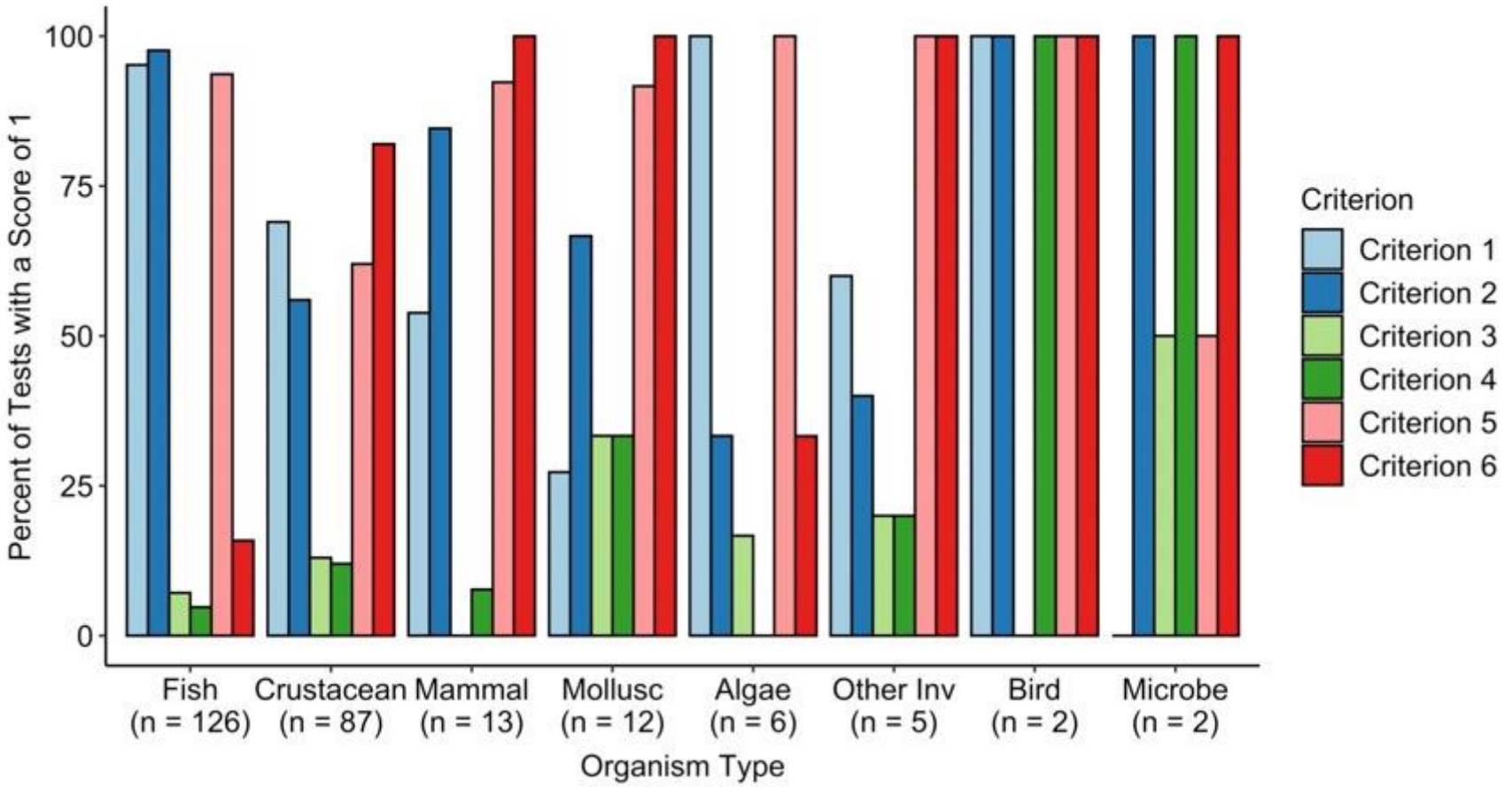

Figure S11 Percentage of unique toxicity tests (*n* = 253 as of June 2021) that achieved a score of 1 for each criterion in Group A (test substance). Criterion 1: test substance purity and/or grade reported; Criterion 2: any measured concentrations analytically confirmed; Criterion 3: individual initial measured concentrations analytically confirmed; Criterion 4: individual final measured concentrations analytically confirmed; Criterion 5: greater than or equal to three concentrations tested, excluding control; Criterion 6: at least one ecologically relevant concentration tested.

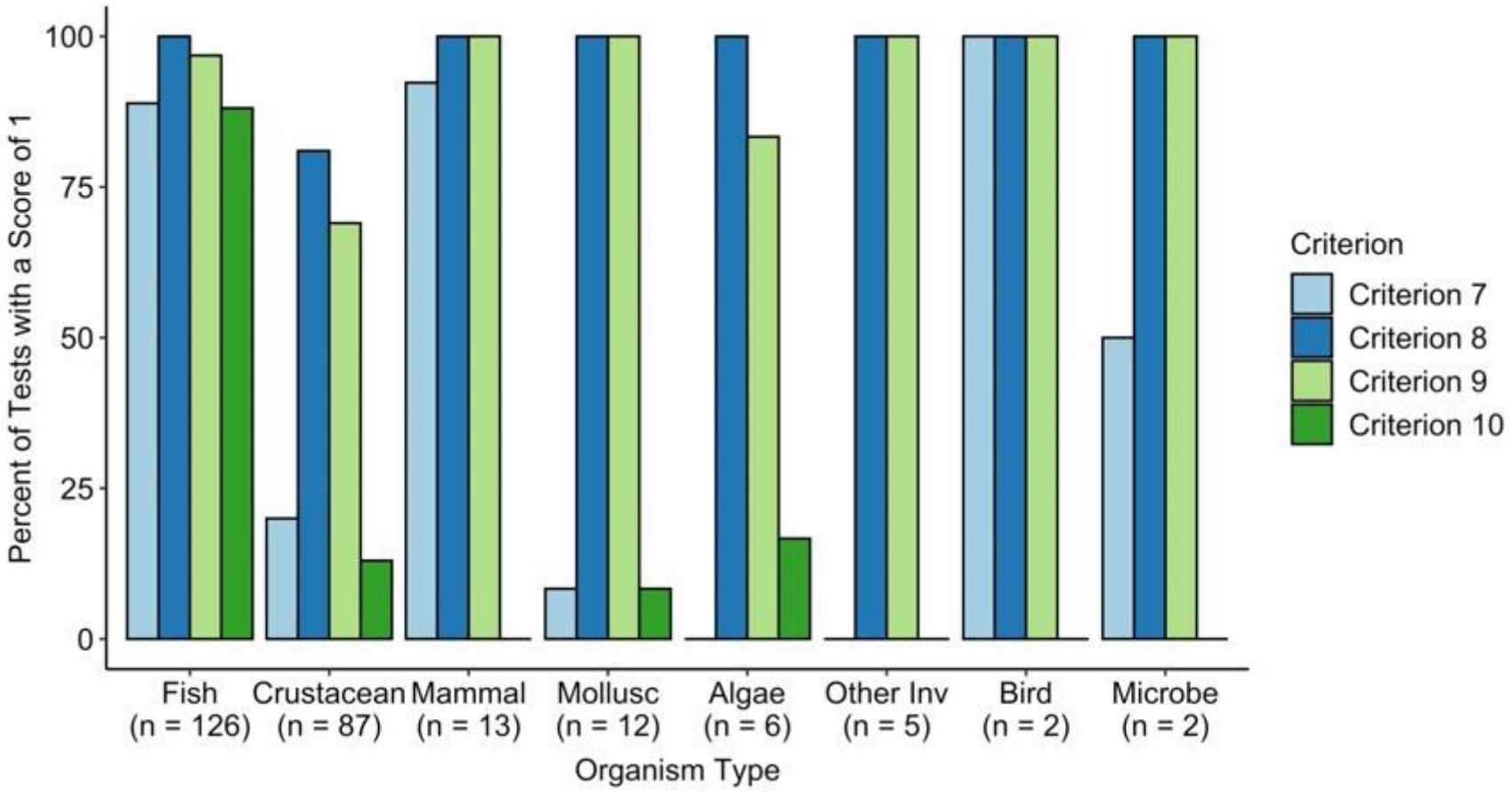

Figure S12 Percentage of unique toxicity tests (*n* = 253 as of June 2021) that achieved a score of 1 for each criterion in Group B (test organism and test system). Criterion 7: strain or source identified; Criterion 8: initial test organism characteristics described; Criterion 9: standard protocol followed; Criterion 10: test conditions reported.

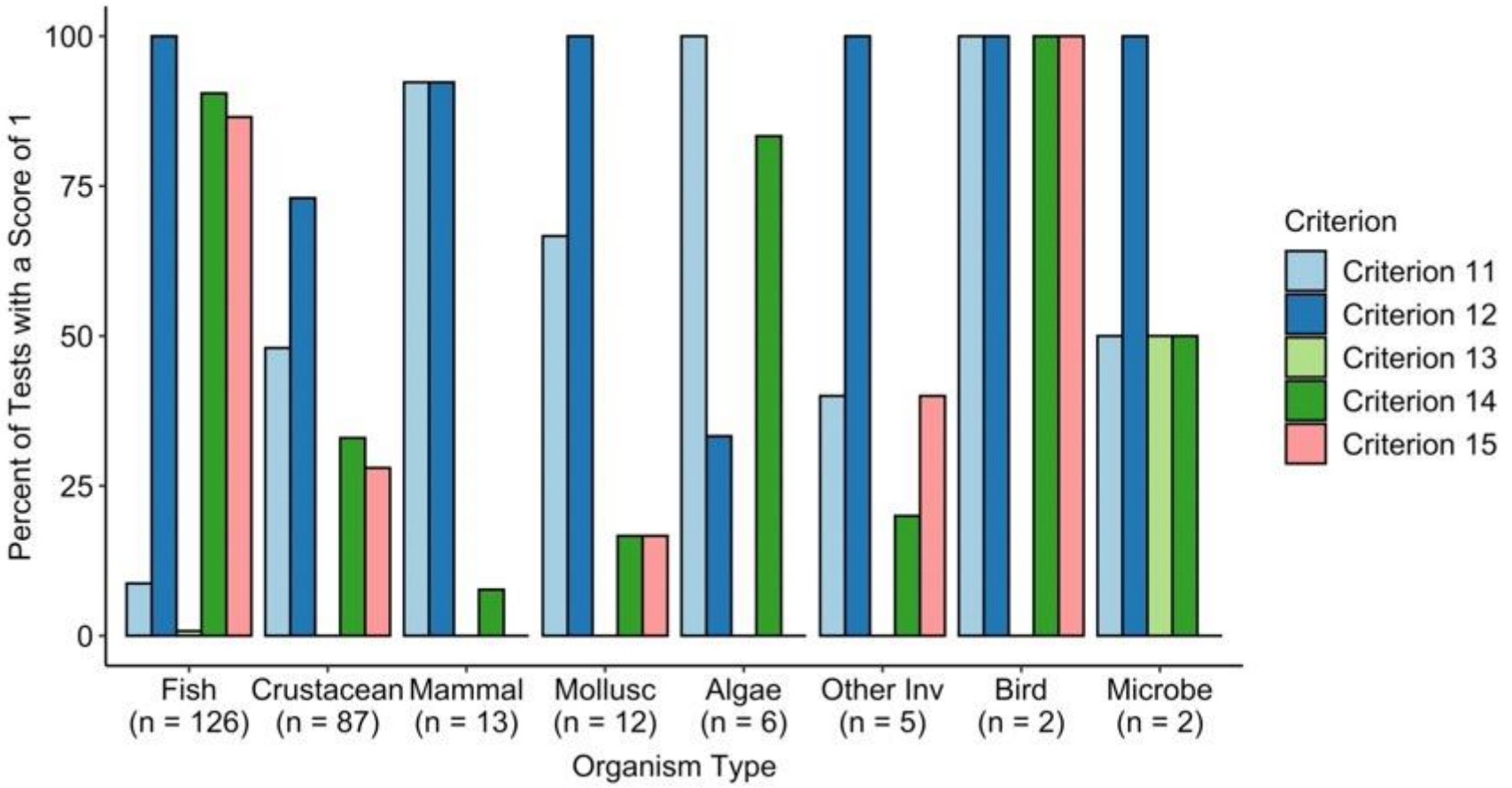

Figure S13 Percentage of unique toxicity tests (*n* = 253 as of June 2021) that achieved a score of 1 for each criterion in Group C (test design, statistics, and results). Criterion 11: greater than or equal to three replicates included; Criterion 12: statistical methods described; Criterion 13: concentration-response model and parameters provided; Criterion 14: raw data reported; Criterion 15: control values reported and criteria met.

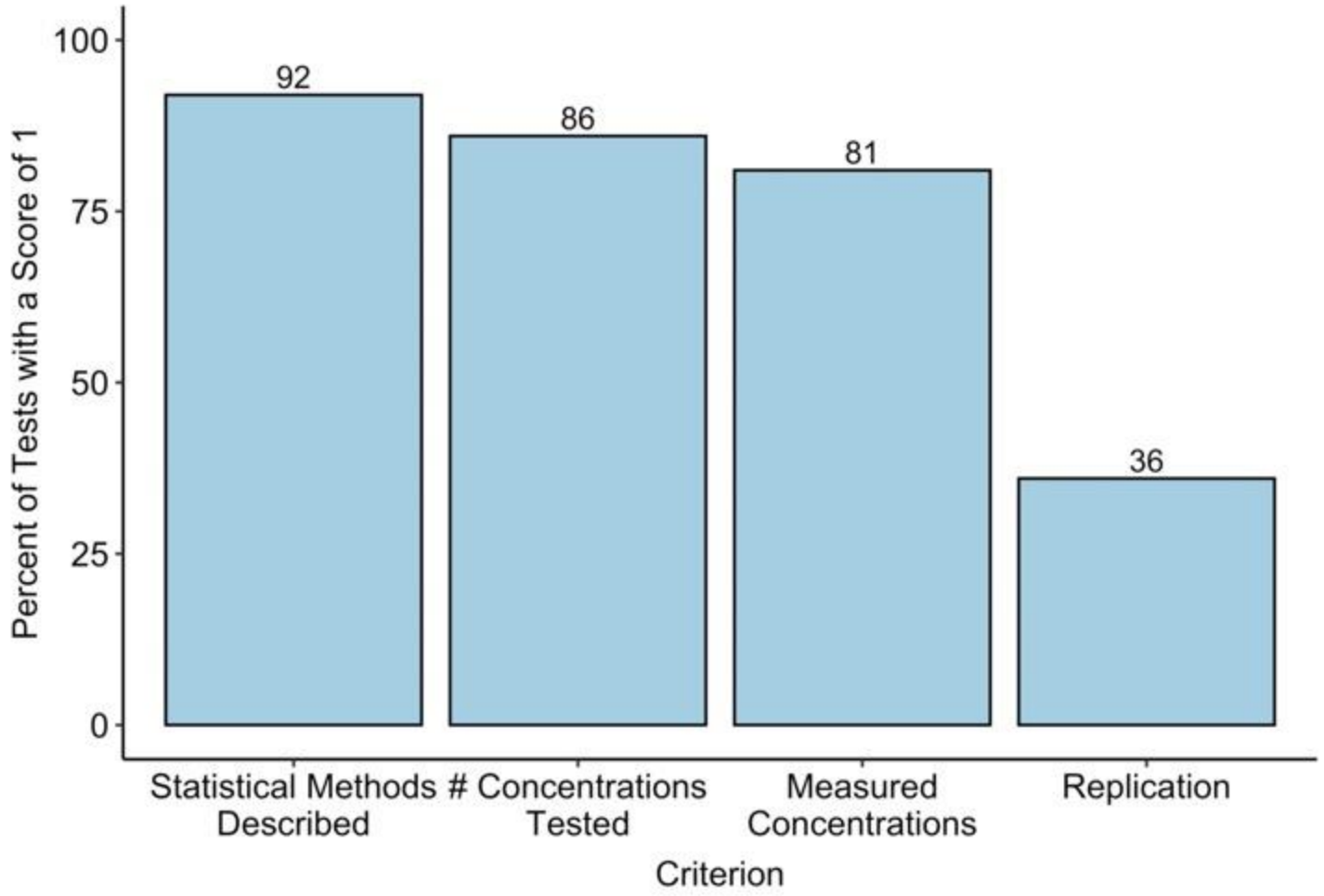

Figure S14 Percentage of unique toxicity tests (*n* = 253) that achieved a score of 1 in each critical study criterion outlined in the scoring rubric presented in this study. Number above each bar corresponds to percentage of tests with a score of 1 for that criterion.

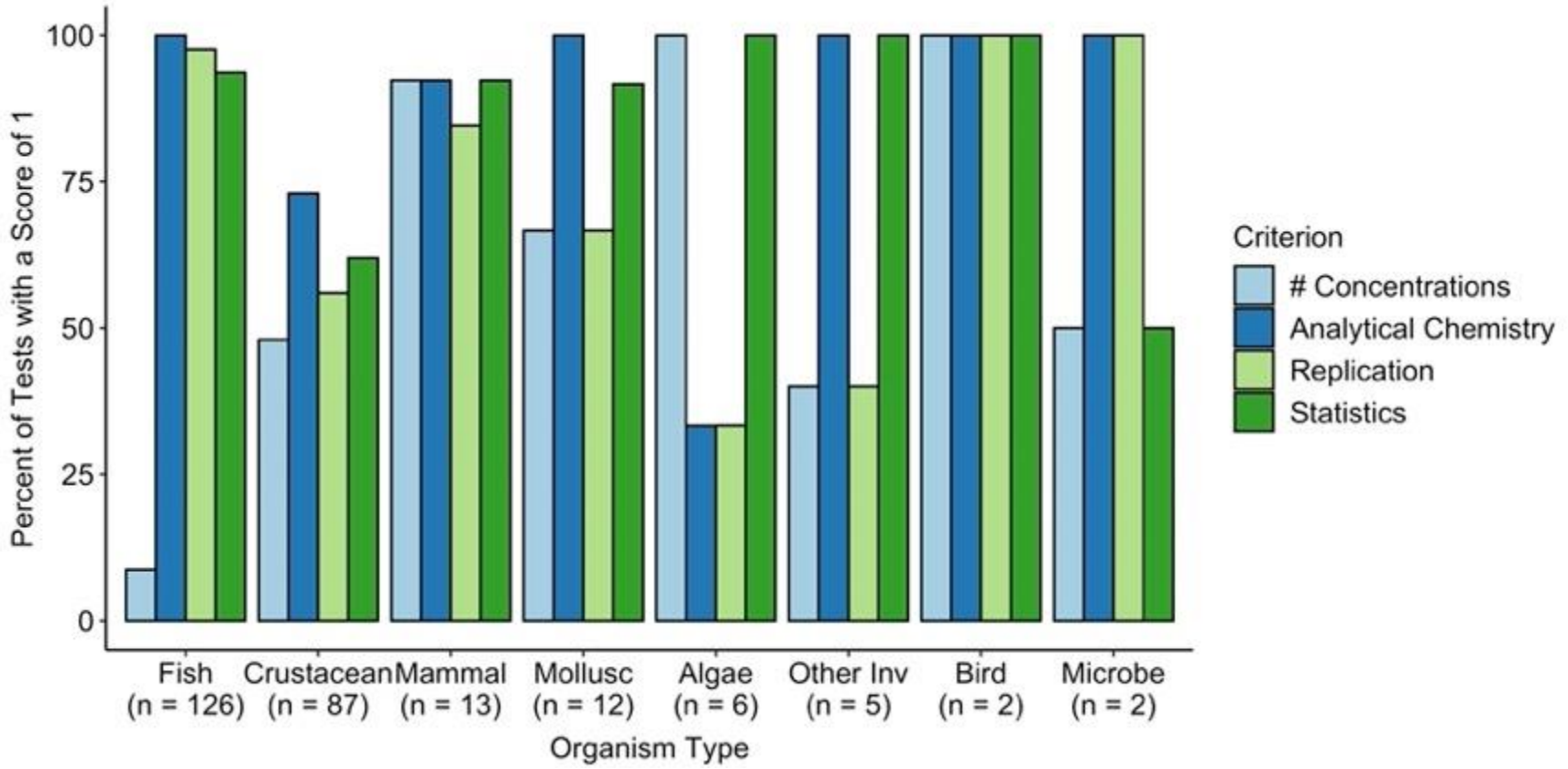

Figure S15 Percentage of unique toxicity tests (*n* = 253) by organism type that achieved a score of 1 in each critical study criterion outlined in the scoring rubric presented in this study.

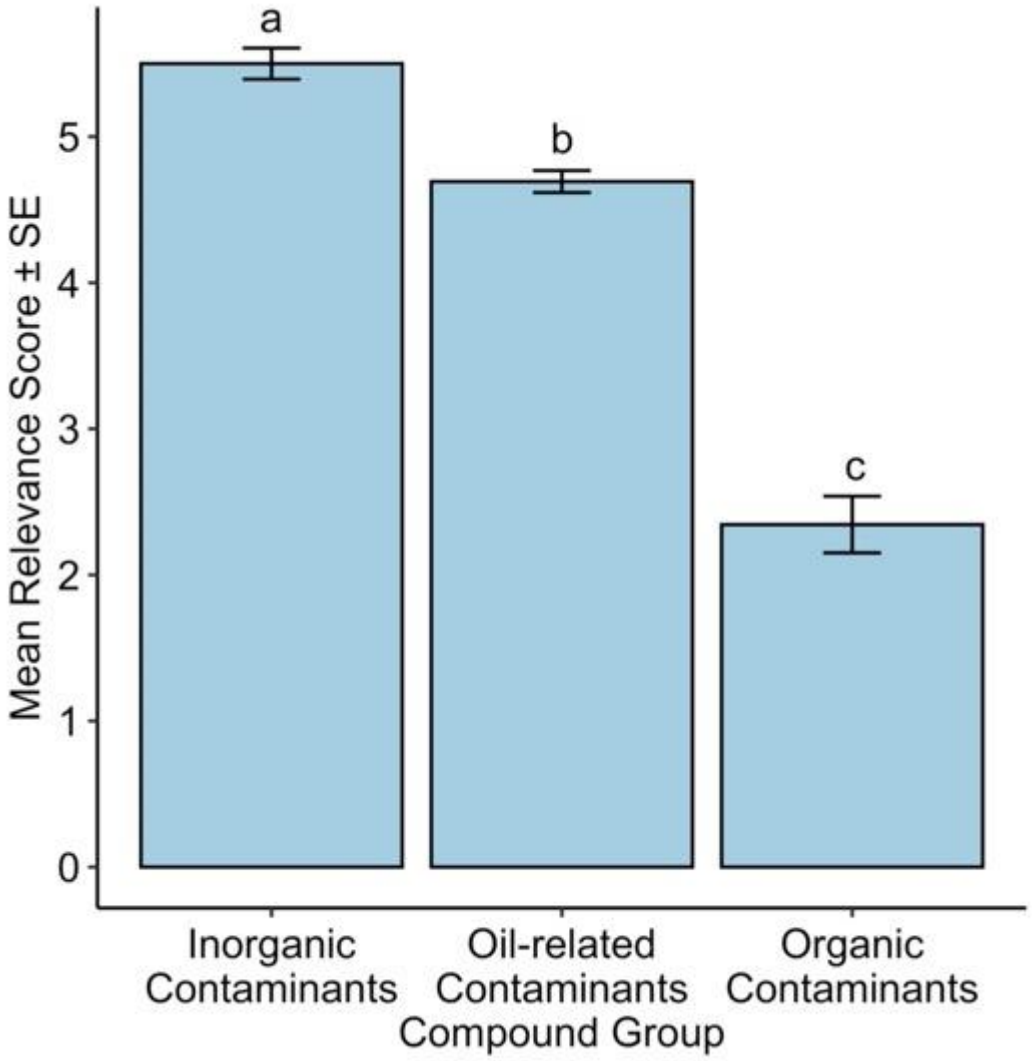

Figure S16 Relevance scores for all experimental combinations assessed in this study (i.e., test substance/test species/endpoint combinations, *n* = 596) by test substance type. Bars are standard error and letters above each column indicate statistical significance between groups ($p < 0.05$) based on Kruskal-Wallis rank sum tests.

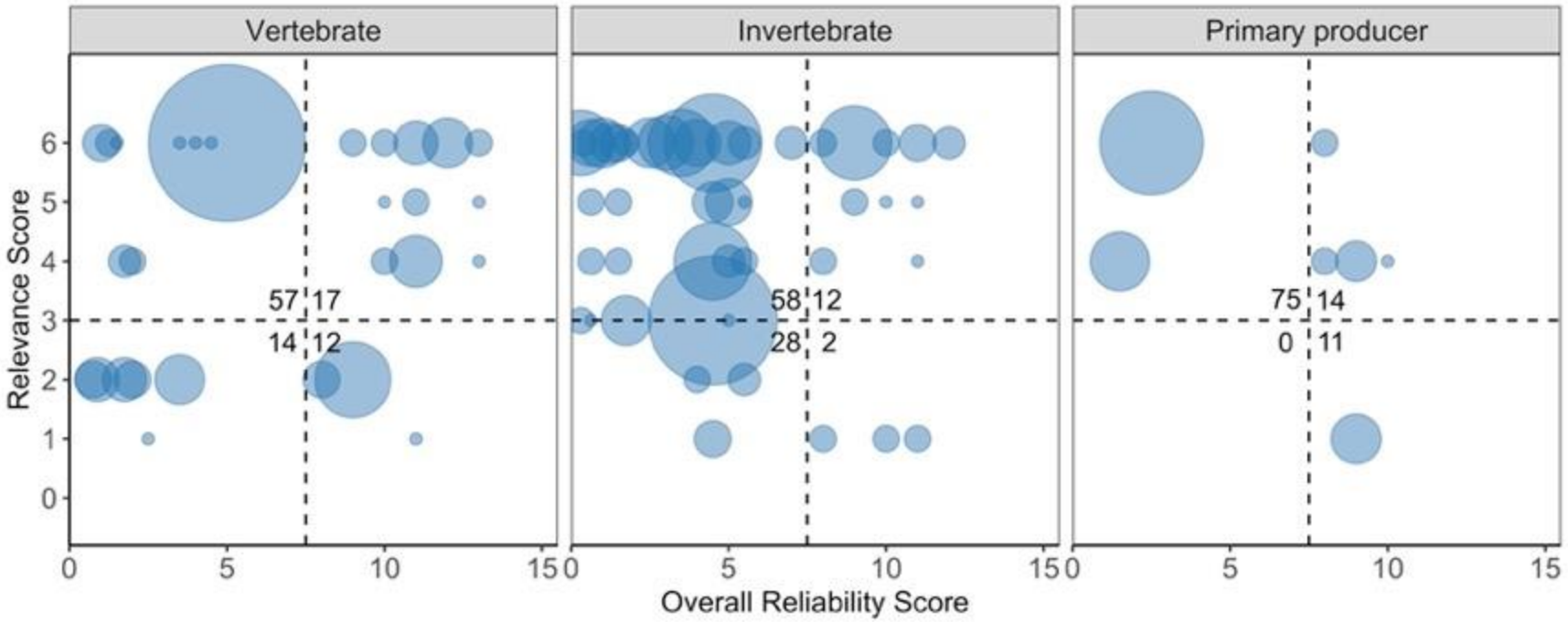

Figure S17 Overall and relevance scores for all experimental combinations collected in this study by organism group (*n* = 596). Size of circle corresponds to number of studies. Numerical values in each quadrant refer to percentage of data points with a relevance score between 0 and 3 (inclusive) and an overall score of < 50% (bottom left quadrant) or > 50% (bottom right quadrant), and relevance score between 4 and 6 (inclusive) and an overall score of < 50% (top left quadrant) or > 50% (top right quadrant).

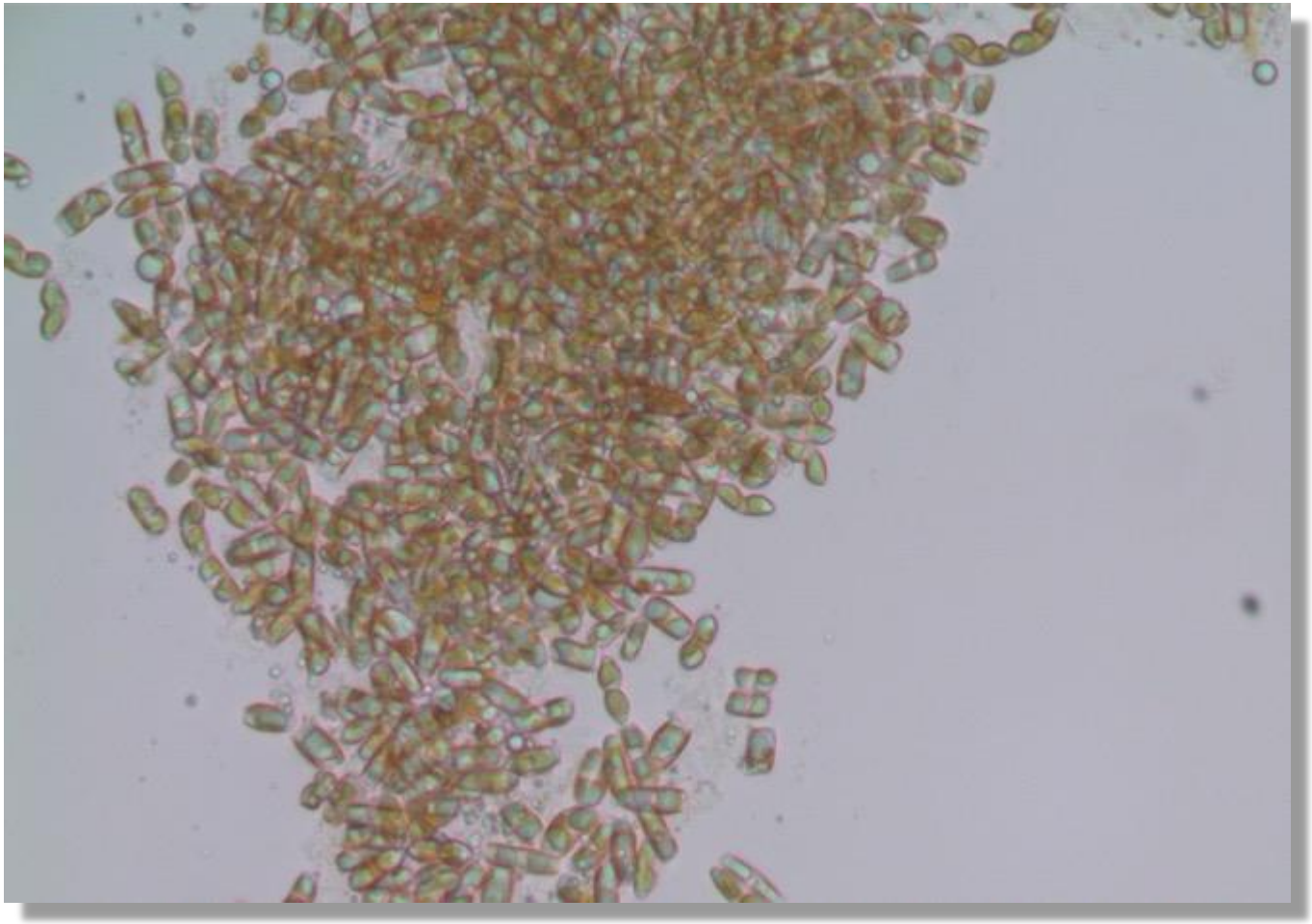

Figure S18 *Nitzschia frigida* cells growing in an arborescent colony (20x magnification).

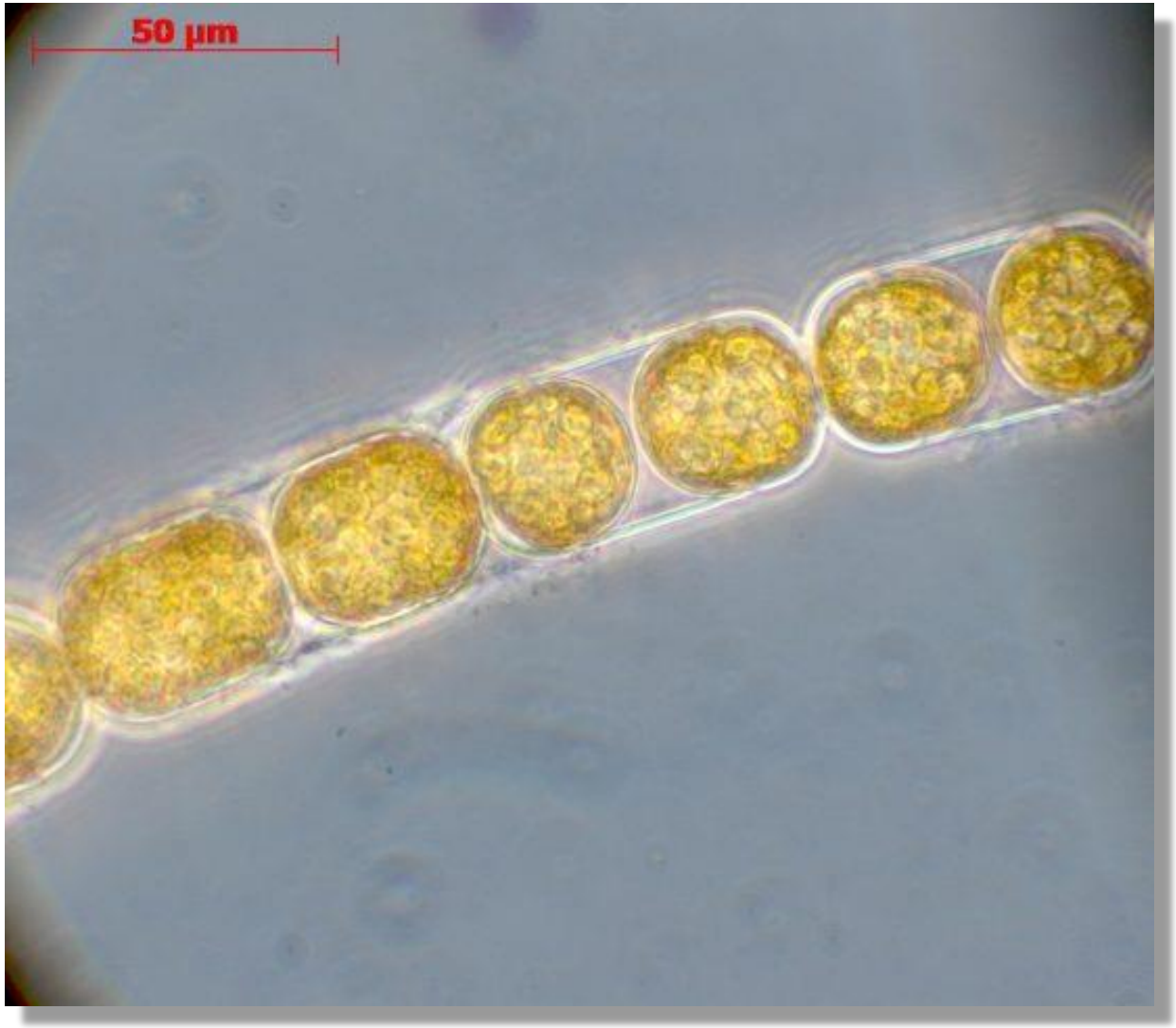

Figure S19 *Melosira arctica* cells growing in filaments (Mares, 2017). Scale bar is 50 µm.

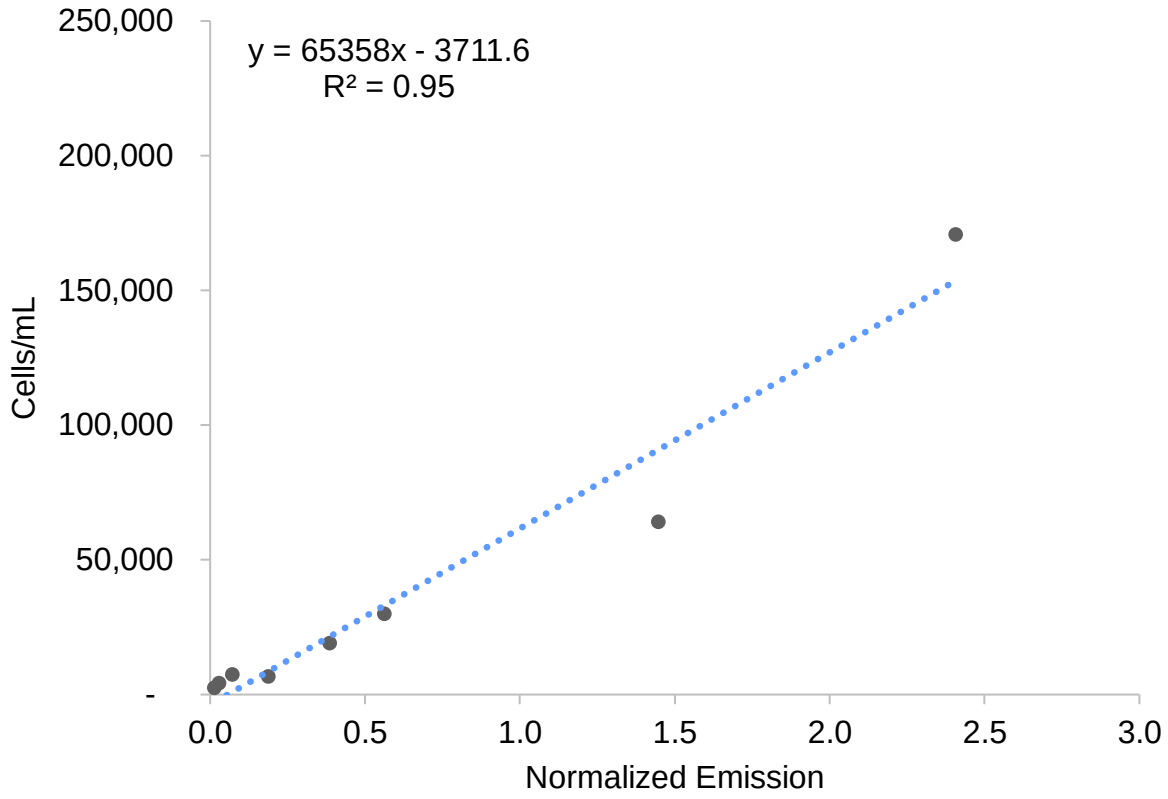

Figure S20 Relationship between manual cell counts using a hemocytometer measured in cells/mL and the normalized emission (i.e., RFU divided by Raman scattering peak collected on the same day) from fluorometry measured in RFU for *Nitzschia frigida* (Standard Curve 1).

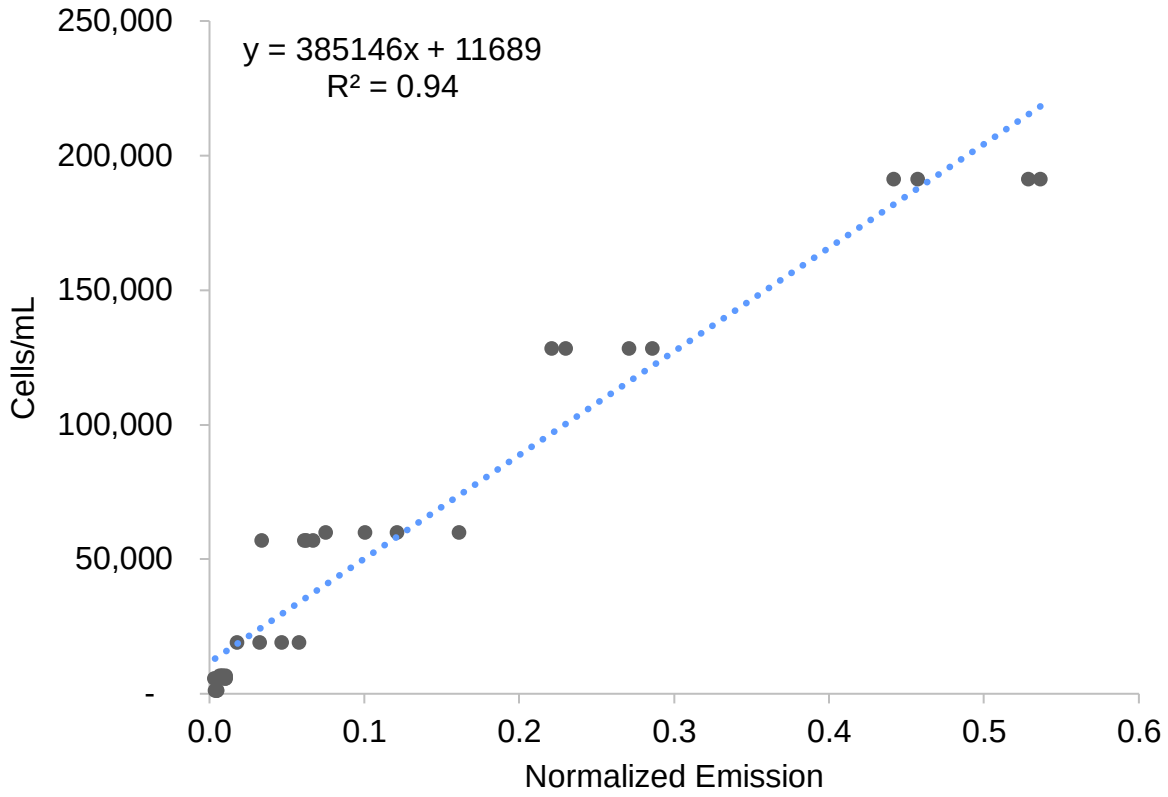

Figure S21 Relationship between manual cell counts using a hemocytometer measured in cells/mL and the normalized emission from fluorometry measured in RFU for *Nitzschia frigida* (Standard Curve 2).

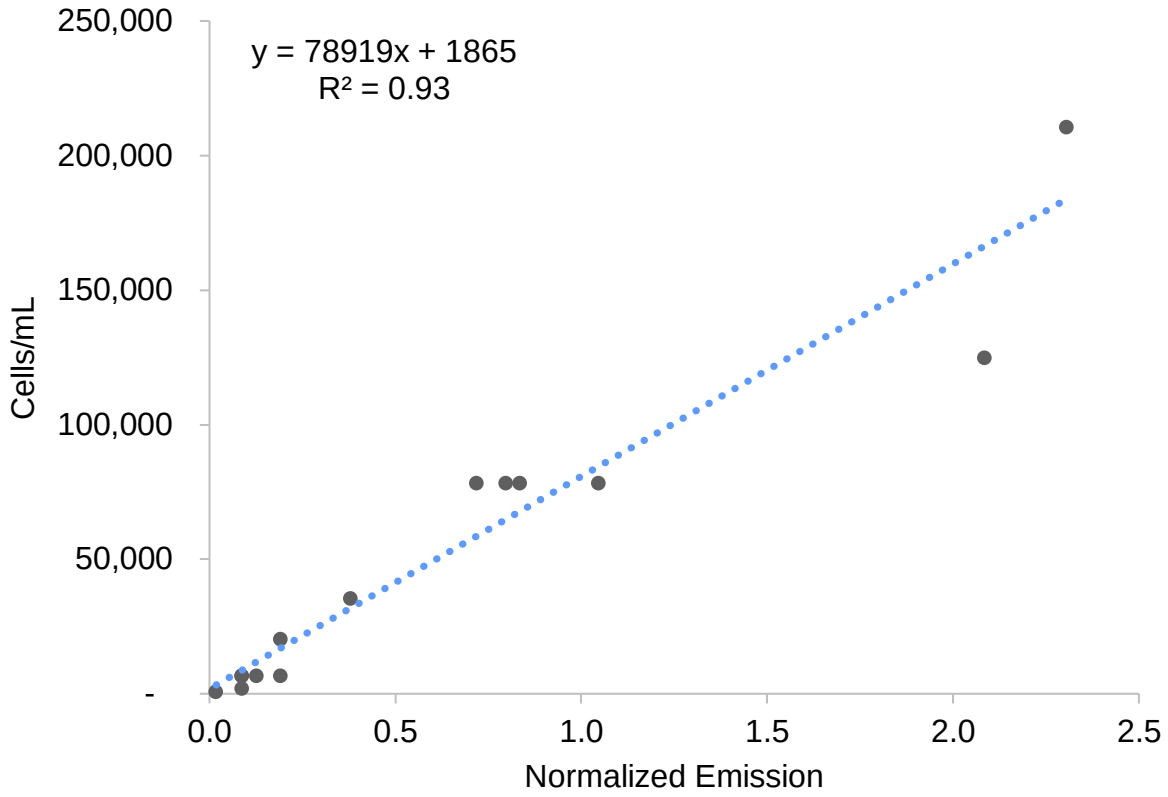

Figure S22 Relationship between manual cell counts using a hemocytometer measured in cells/mL and the normalized emission from fluorometry measured in RFU for *Nitzschia frigida* (Standard Curve 3).

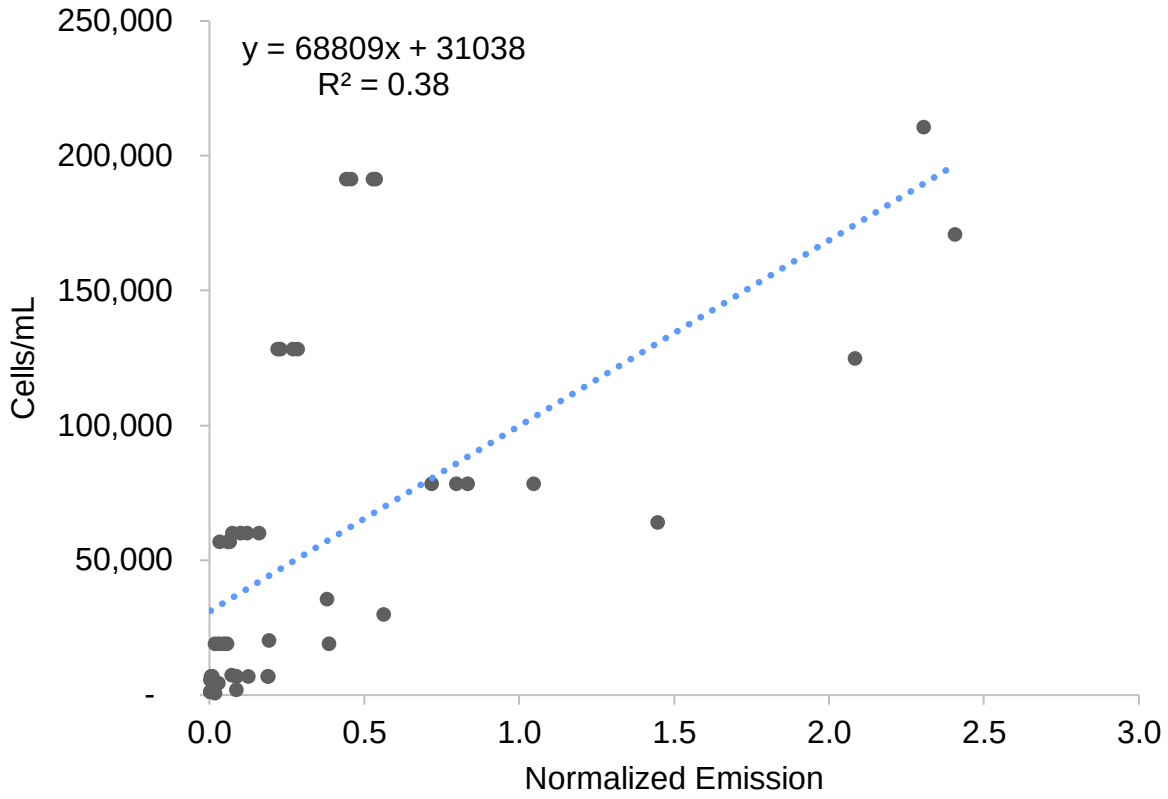

Figure S23 Relationship between manual cell counts using a hemocytometer measured in cells/mL and the normalized emission from fluorometry measured in RFU for *Nitzschia frigida* (Standard Curves 1, 2, and 3 combined).

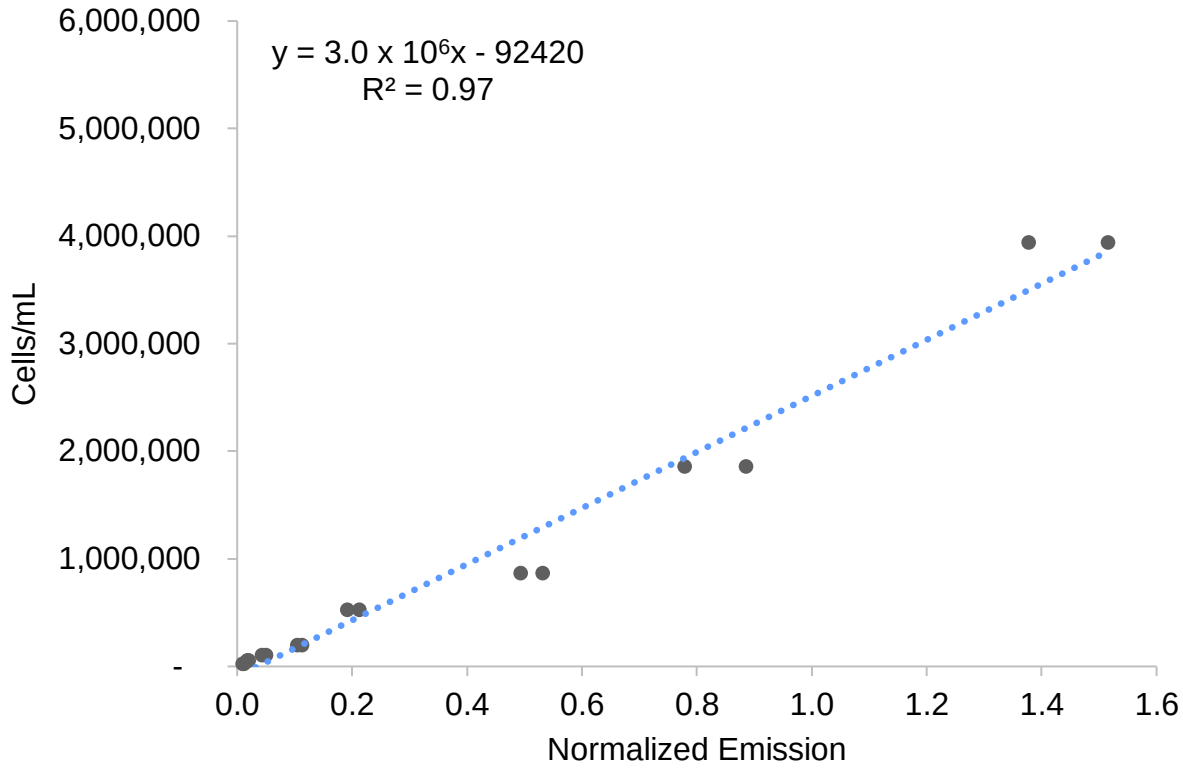

Figure S24 Relationship between manual cell counts using a hemocytometer measured in cells/mL and the normalized emission from fluorometry measured in RFU for *Skeletonema costatum* (Standard Curve 1).

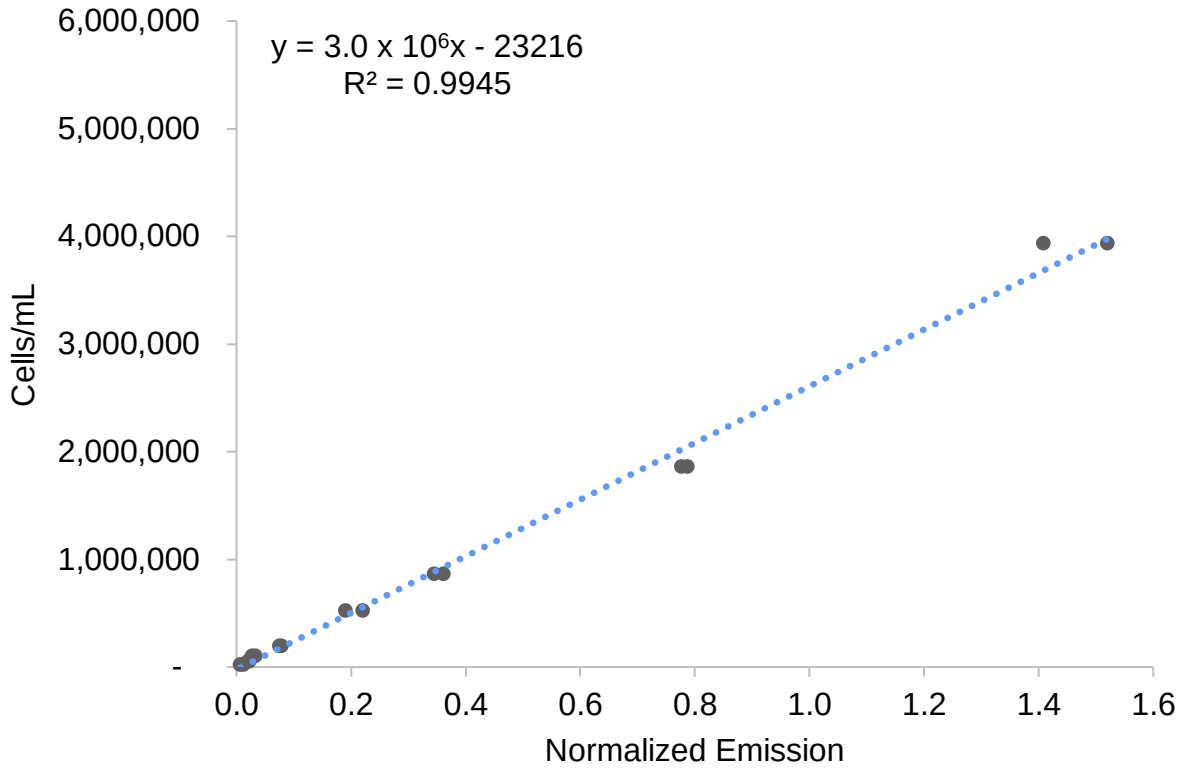

Figure S25 Relationship between manual cell counts using a hemocytometer measured in cells/mL and the normalized emission from fluorometry measured in RFU for *Skeletonema costatum* (Standard Curve 2).

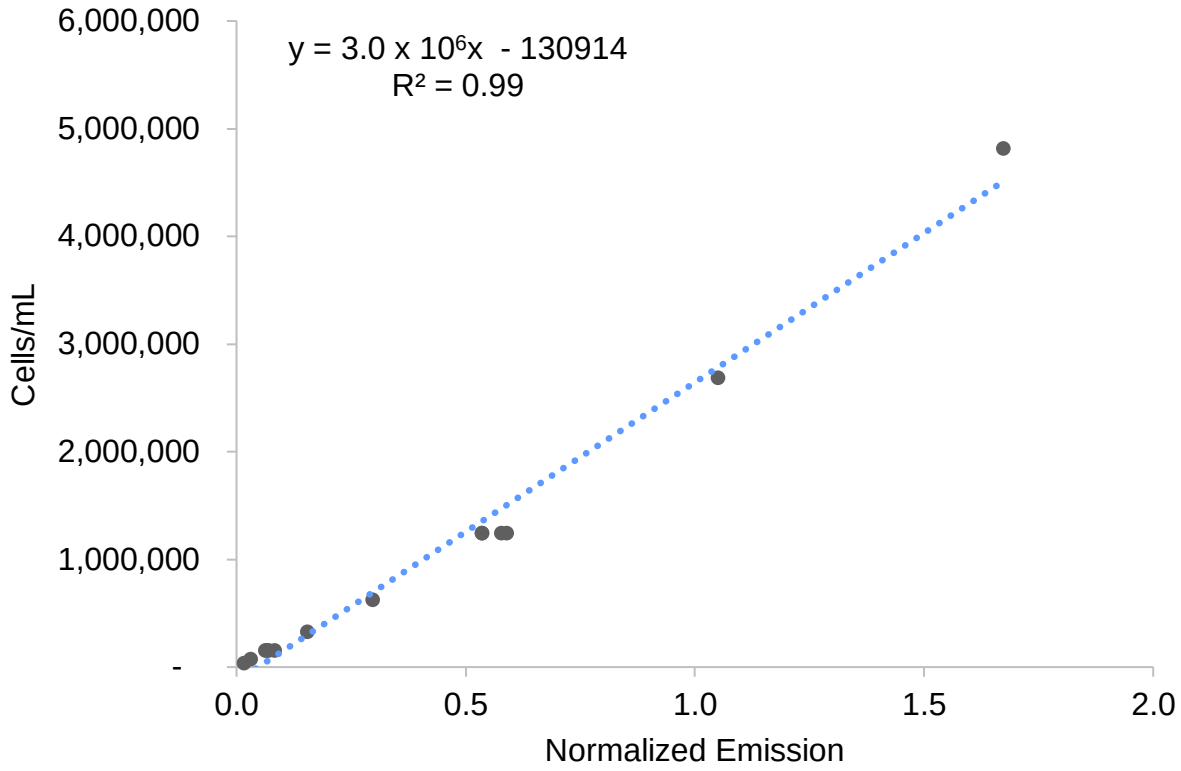

Figure S26 Relationship between manual cell counts using a hemocytometer measured in cells/mL and the normalized emission from fluorometry measured in RFU for *Skeletonema costatum* (Standard Curve 3).

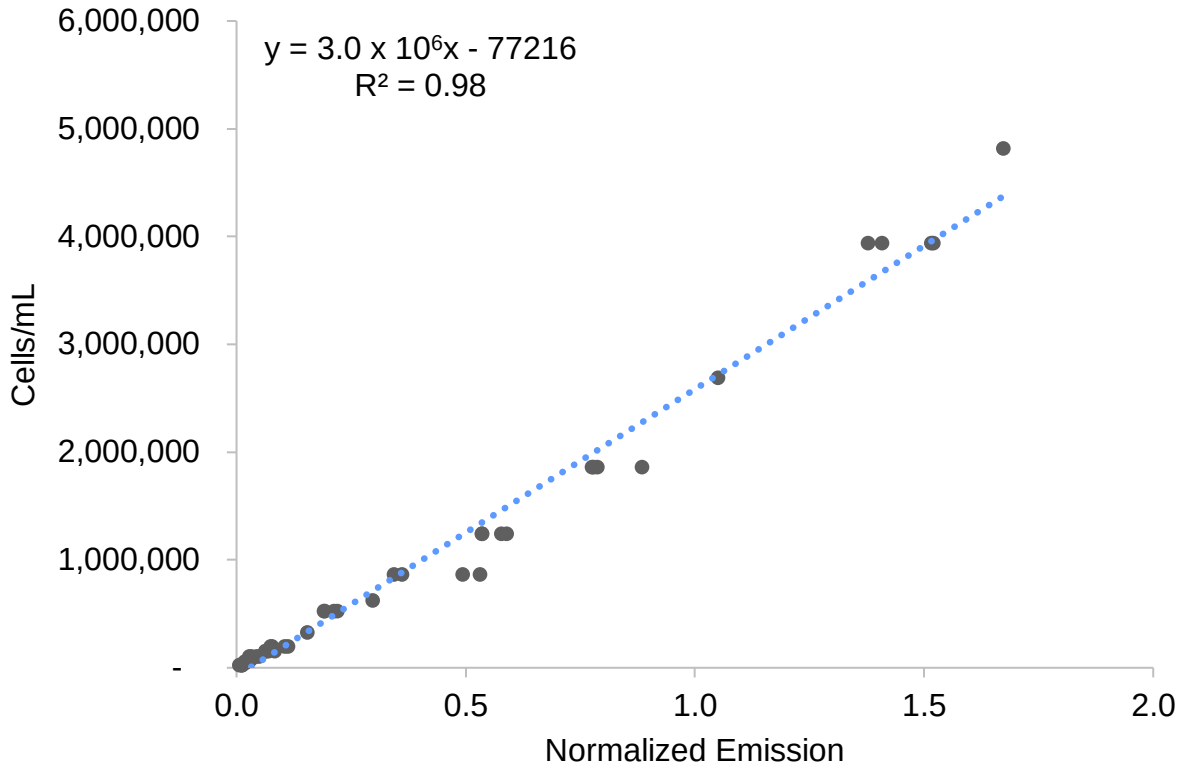

Figure S27 Relationship between manual cell counts using a hemocytometer measured in cells/mL and the normalized emission from fluorometry measured in RFU for *Skeletonema costatum* (Standard Curves 1, 2, and 3 combined).

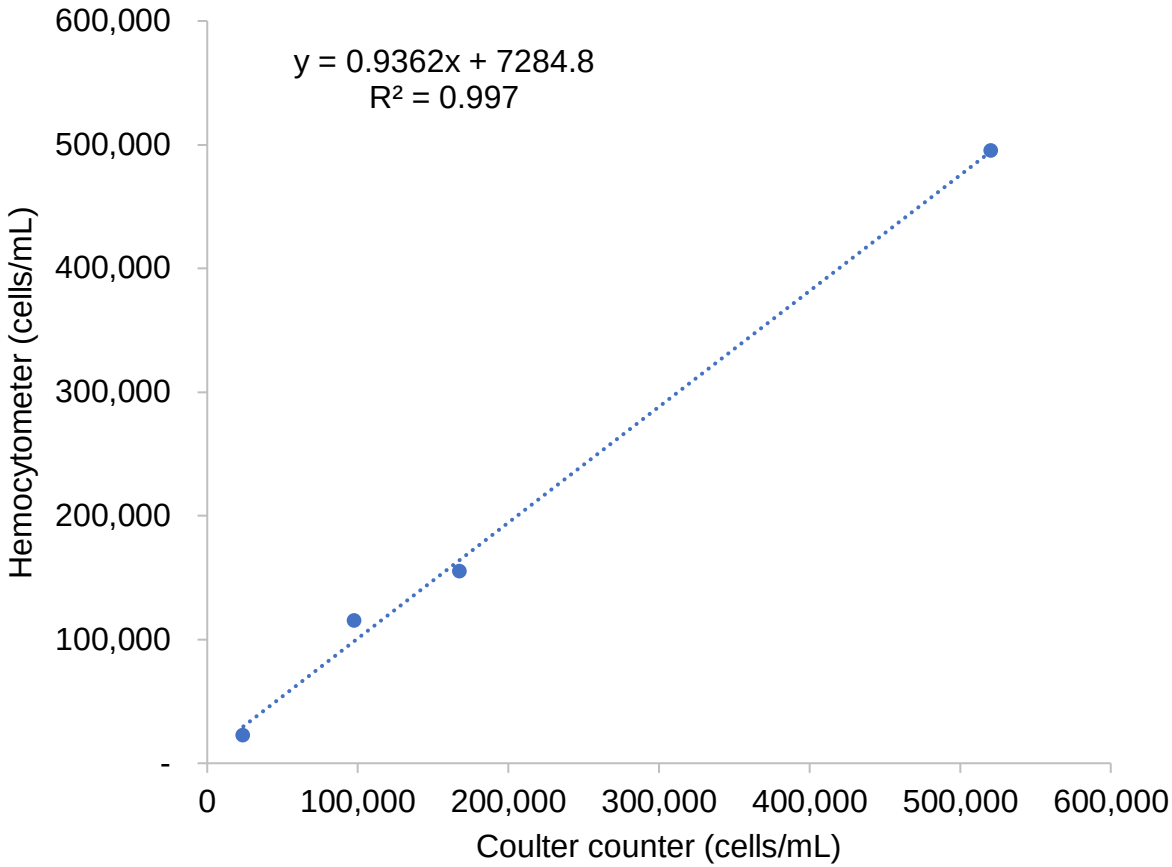

Figure S28 Relationship between manual cell counts using a hemocytometer measured in cells/mL and concentration output from Coulter counter measured in cells/mL for *Nitzschia frigida* (Standard Curve 1).

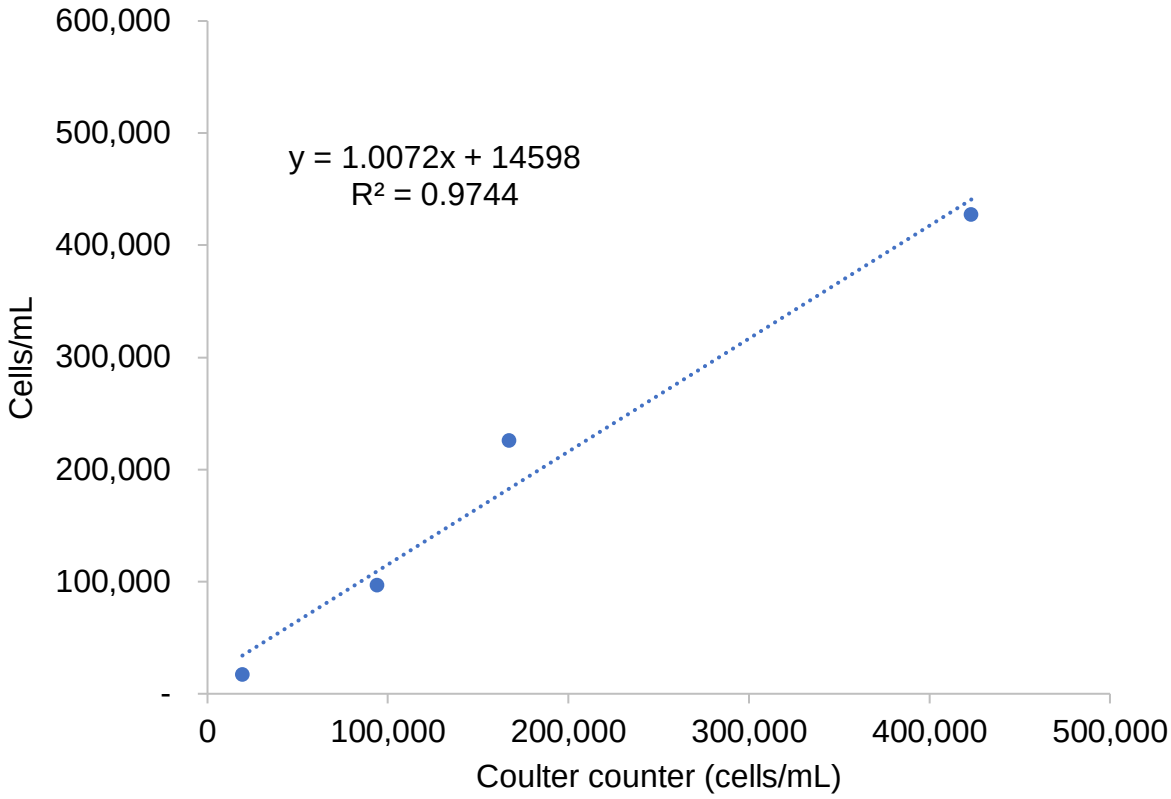

Figure S29 Relationship between manual cell counts using a hemocytometer measured in cells/mL and concentration output from Coulter counter measured in cells/mL for *Nitzschia frigida* (Standard Curve 2).

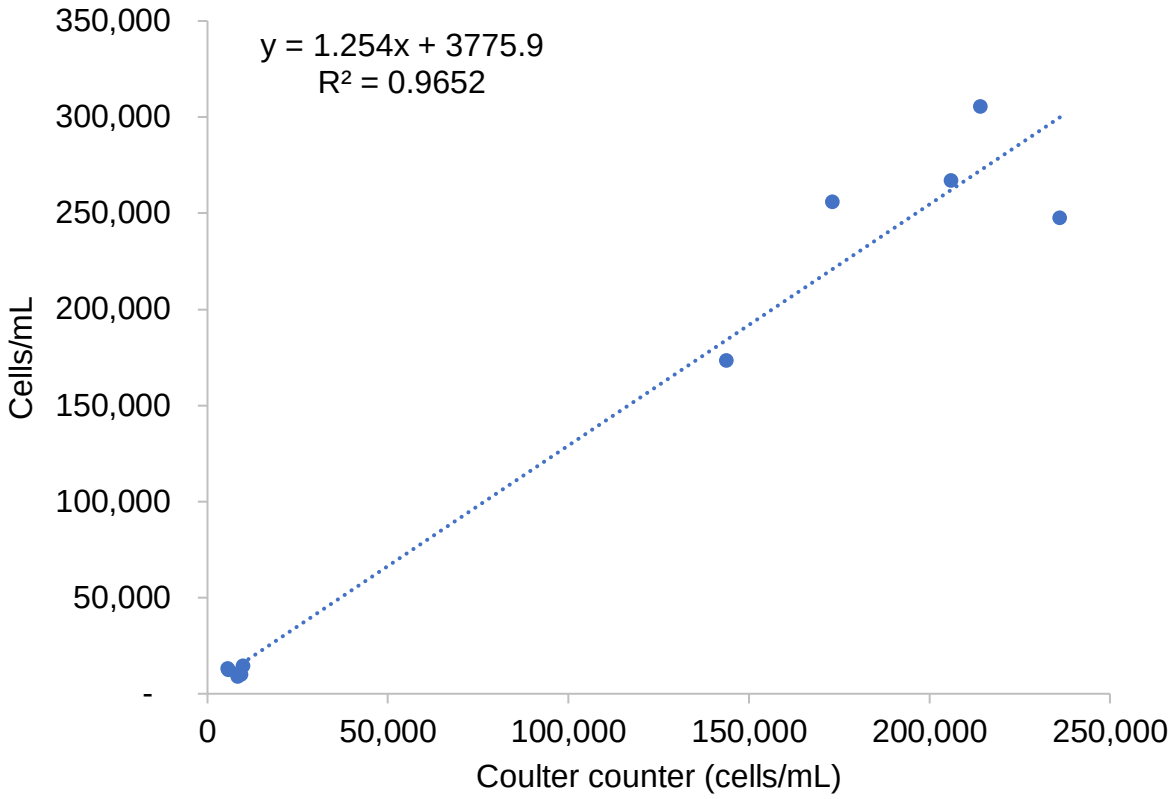

Figure S30 Relationship between manual cell counts using a hemocytometer measured in cells/mL and concentration output from Coulter counter measured in cells/mL for *Nitzschia frigida* (Standard Curve 3).

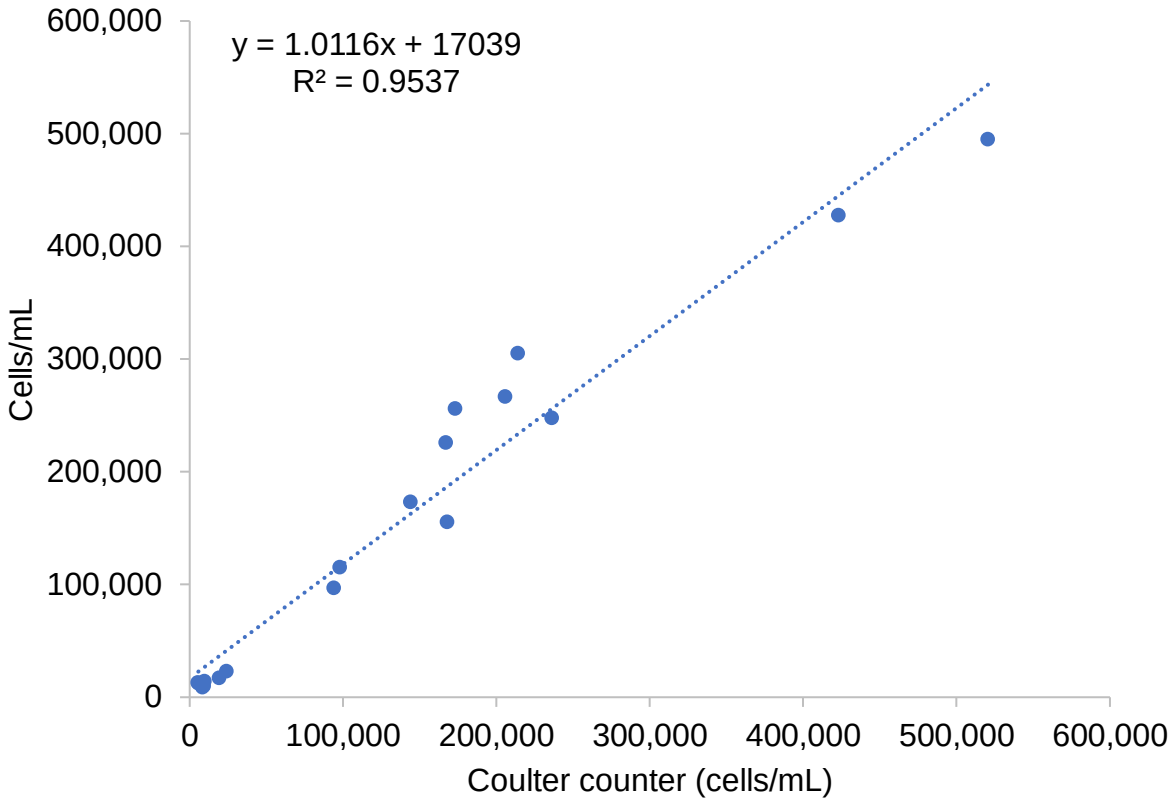

Figure S31 Relationship between manual cell counts using a hemocytometer measured in cells/mL and concentration output from Coulter counter measured in cells/mL for *Nitzschia frigida* (Standard Curve 1, 2, and 3 combined).

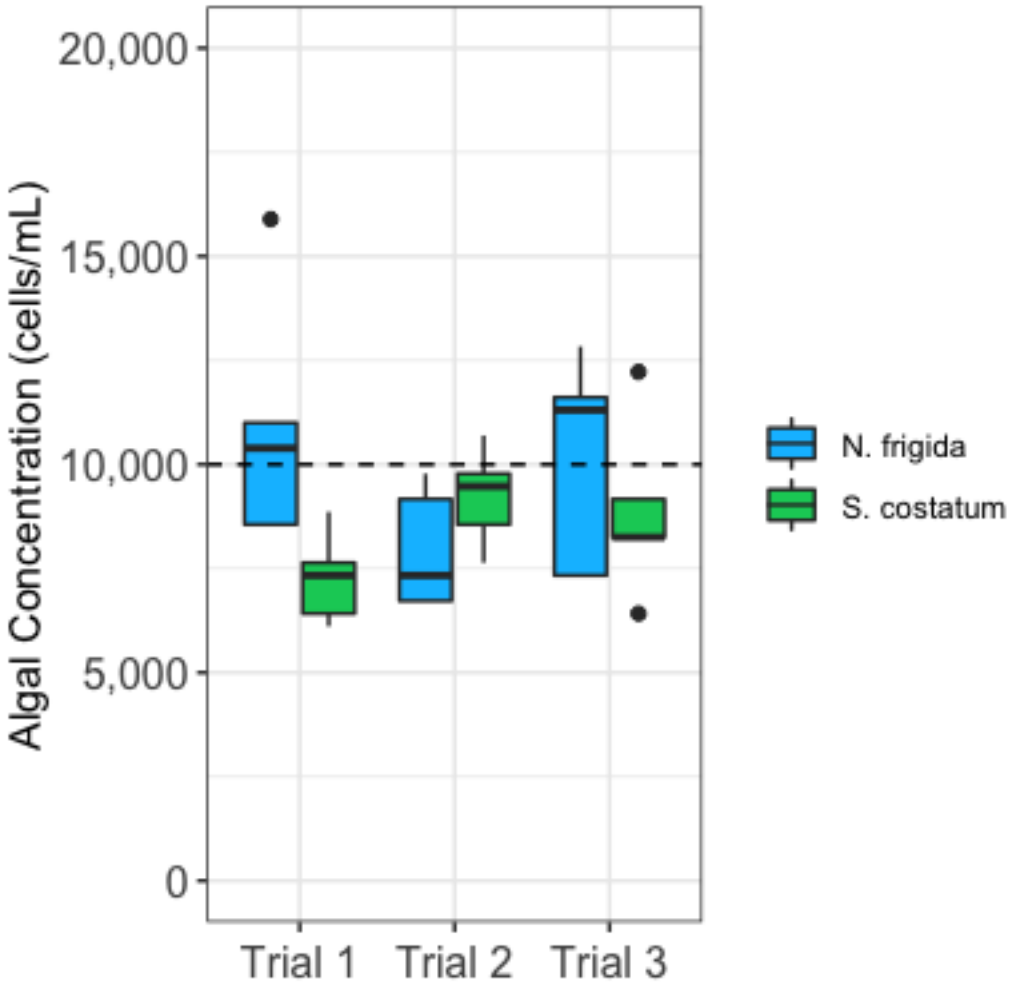

Figure S32 Concentration of Inoculum control units across all three trials conducted in the present study. Dashed line represents target concentration of 10,000 cells/mL. No statistically significant differences ($p < 0.05$) were observed between inoculum control concentrations across trials for either species based on one-way ANOVA followed by Tukey's HSD test.

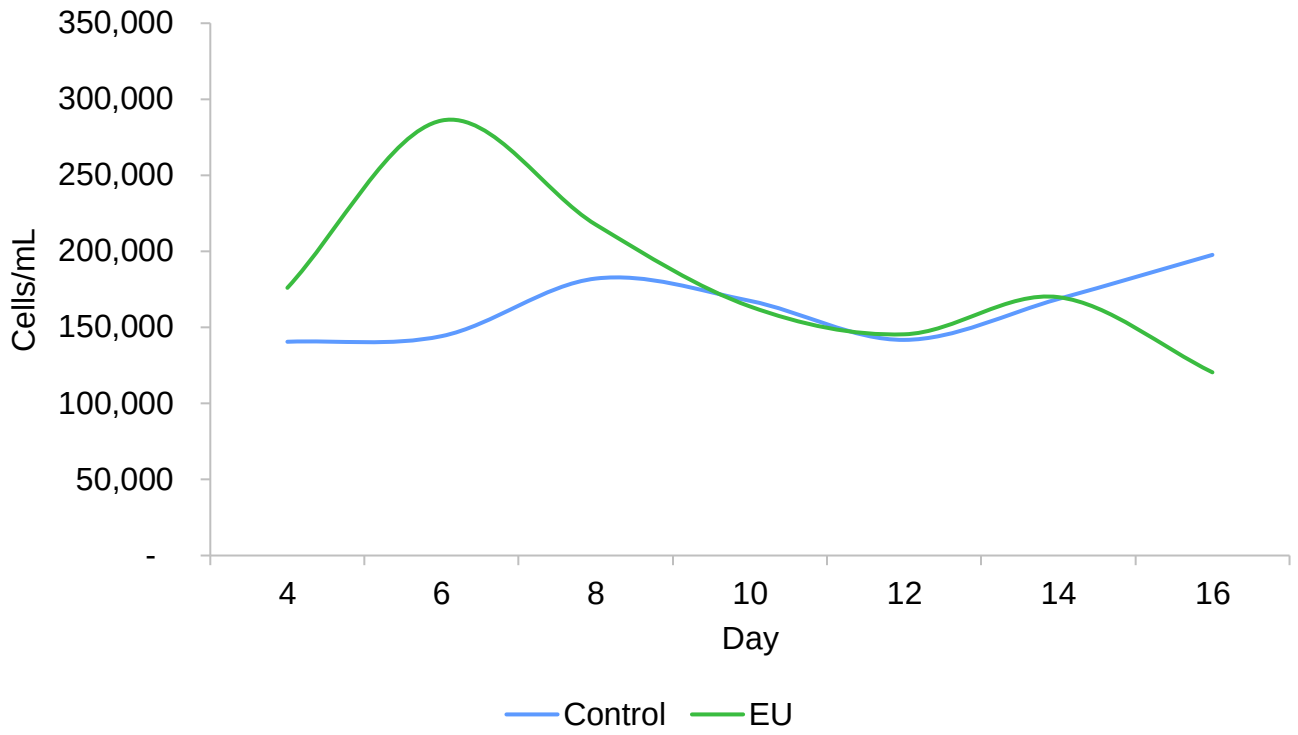

Figure S33 Concentration of algal cells in *Nitzschia frigida* culture (experimental unit; EU) as temperature was increased 2ºC per day from 4ºC to 16ºC. Control unit was monitored at a constant 2ºC throughout the duration as a comparison.

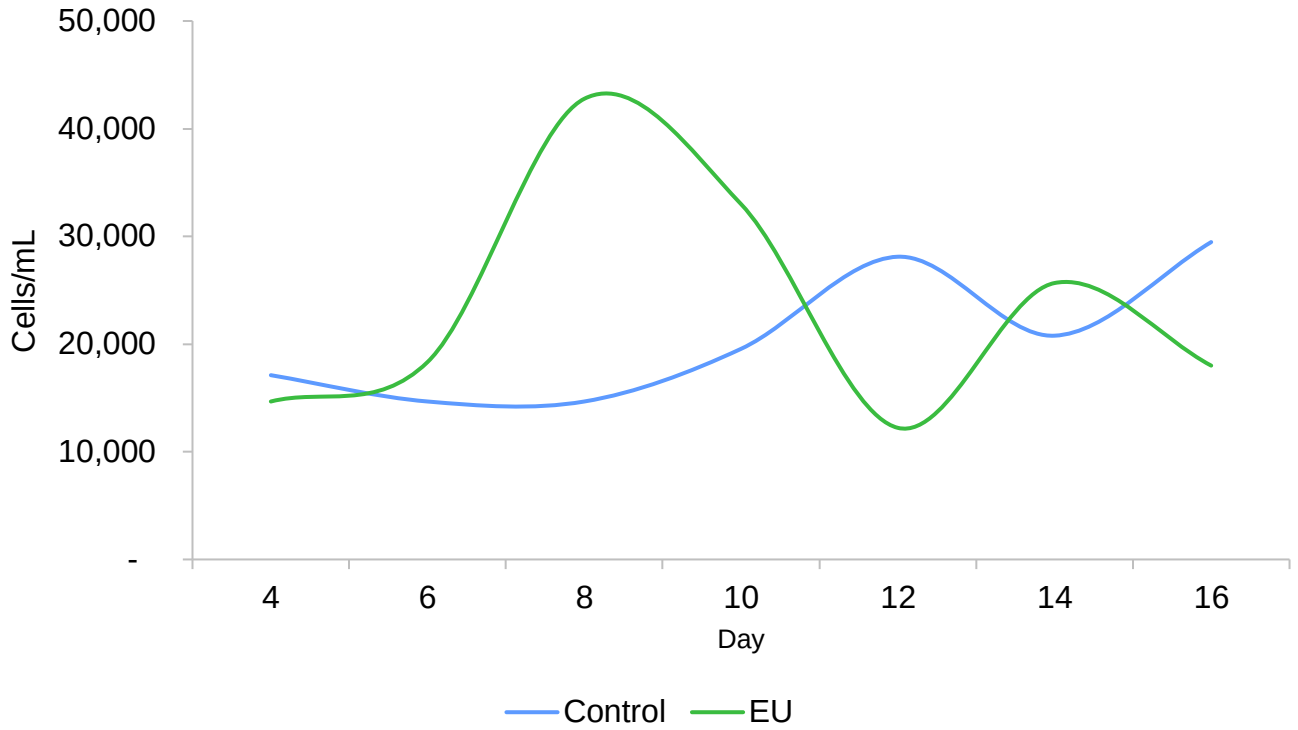

Figure S34 Concentration of algal cells in *Porosira glacialis* culture (experimental unit; EU) as temperature was increased 2ºC per day from 4ºC to 16ºC. Control unit was monitored at a constant 2ºC throughout the duration as a comparison.

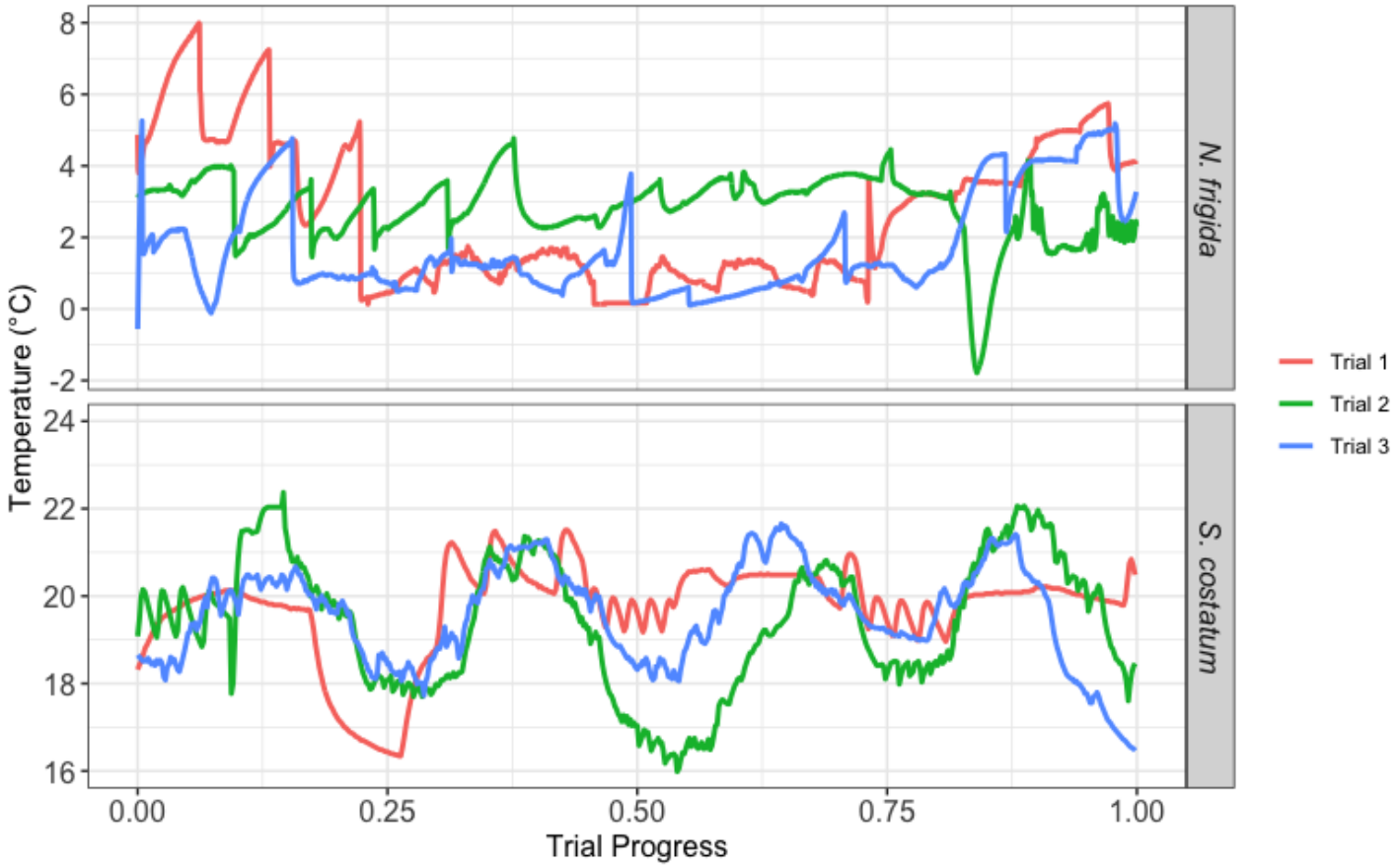

Figure S35 Temperature profiles for all trials from TidbiT datalogger. Trial progress on the x-axis represents fraction of total elapsed time of exposure.

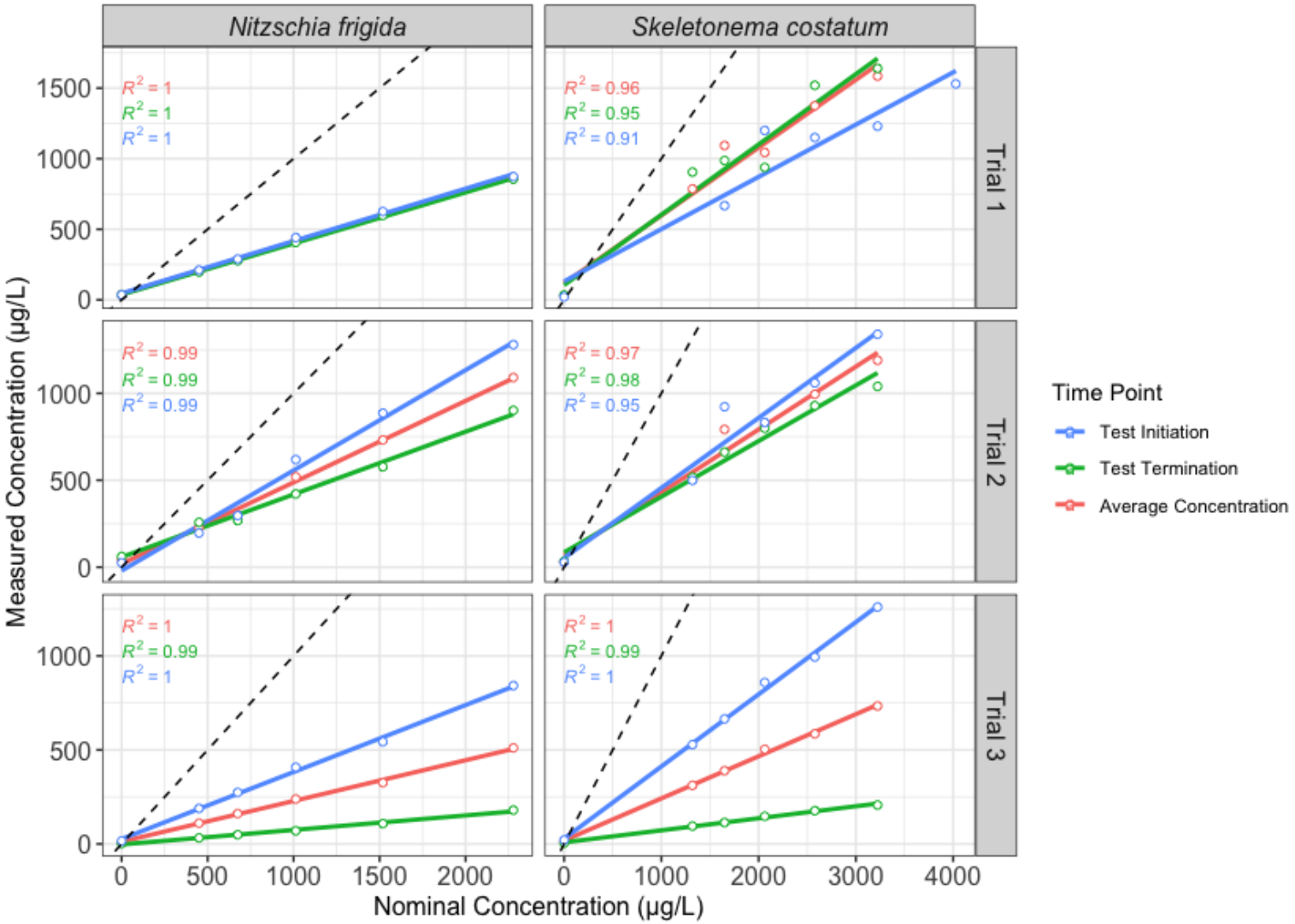

Figure S36 Relationship between measured and nominal copper concentrations for each trial conducted in this study based on analytical chemistry results from a third-party laboratory. Dashed line indicates a 1:1 relationship.

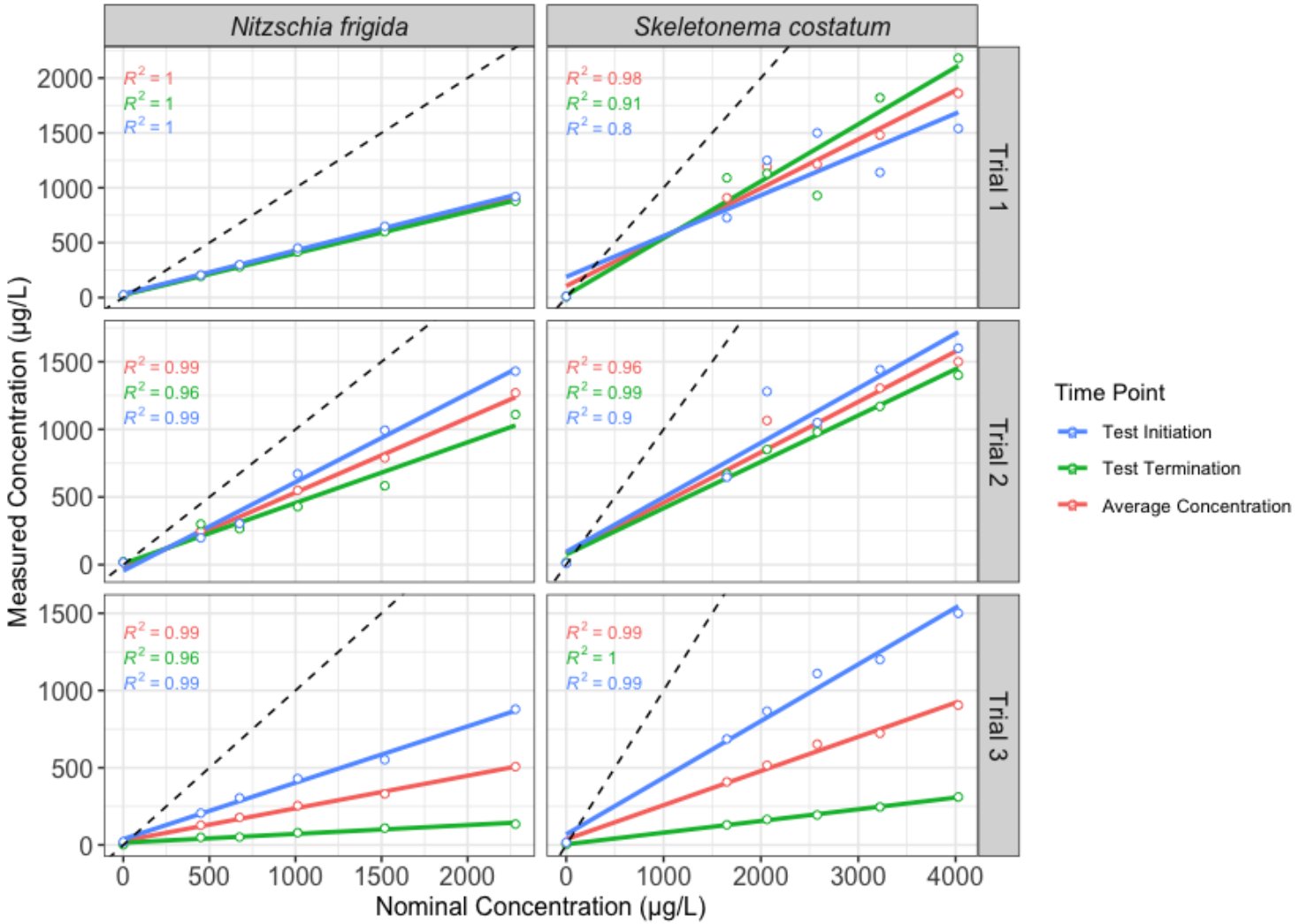

Figure S37 Relationship between measured and nominal zinc concentrations for each trial conducted in this study based on analytical chemistry results from a third-party laboratory. Dashed line indicates a 1:1 relationship.

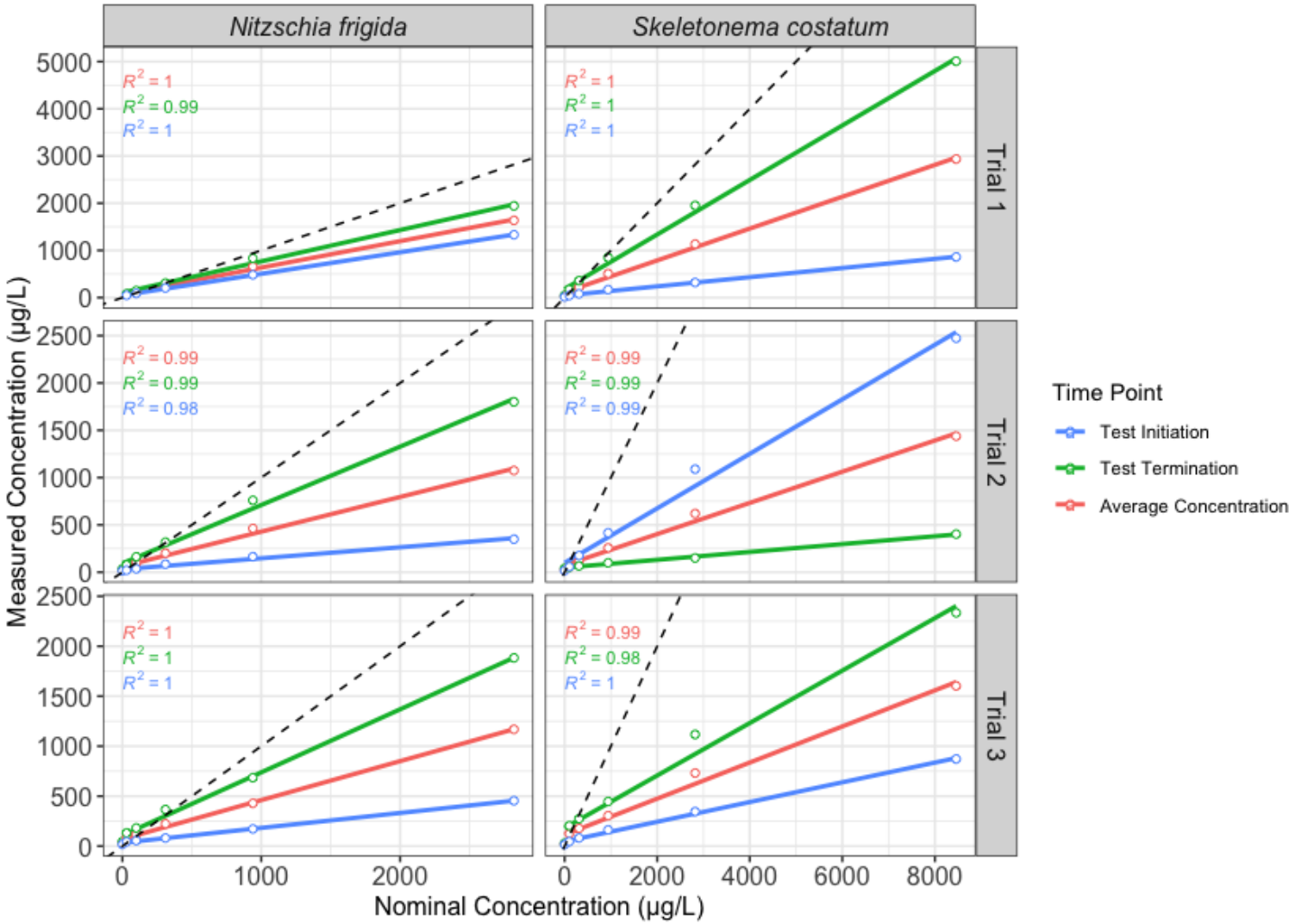

Figure S38 Relationship between measured and nominal methylnaphthalene concentrations for each trial conducted in this study based on in-house analytical chemistry using fluorometry. Dashed line indicates a 1:1 relationship.

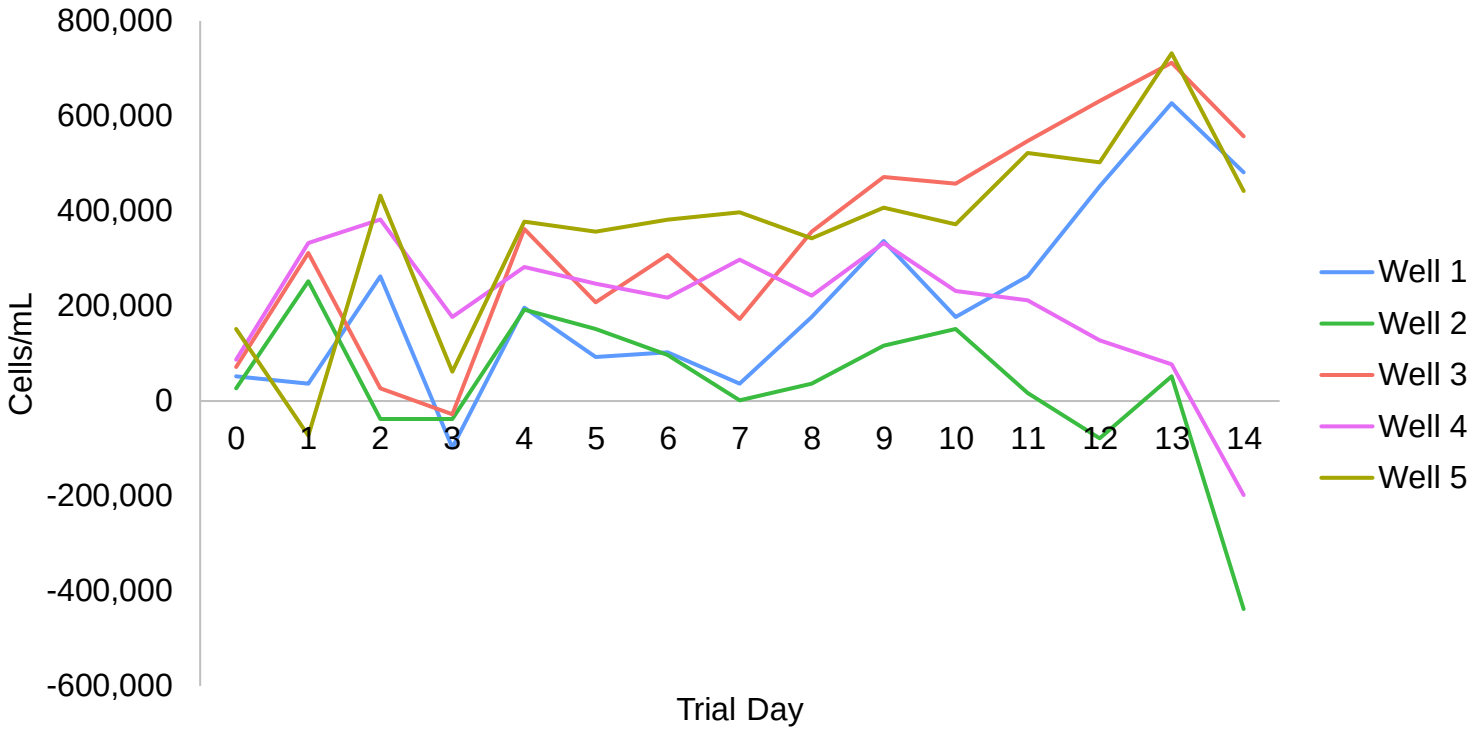

Figure S39 Growth responses of *Porosira glacialis* culture in 96-well microplate under control conditions over 14 days. Measurements from five wells were taken daily using a microplate reader and converted into cells/mL based on standard curves.

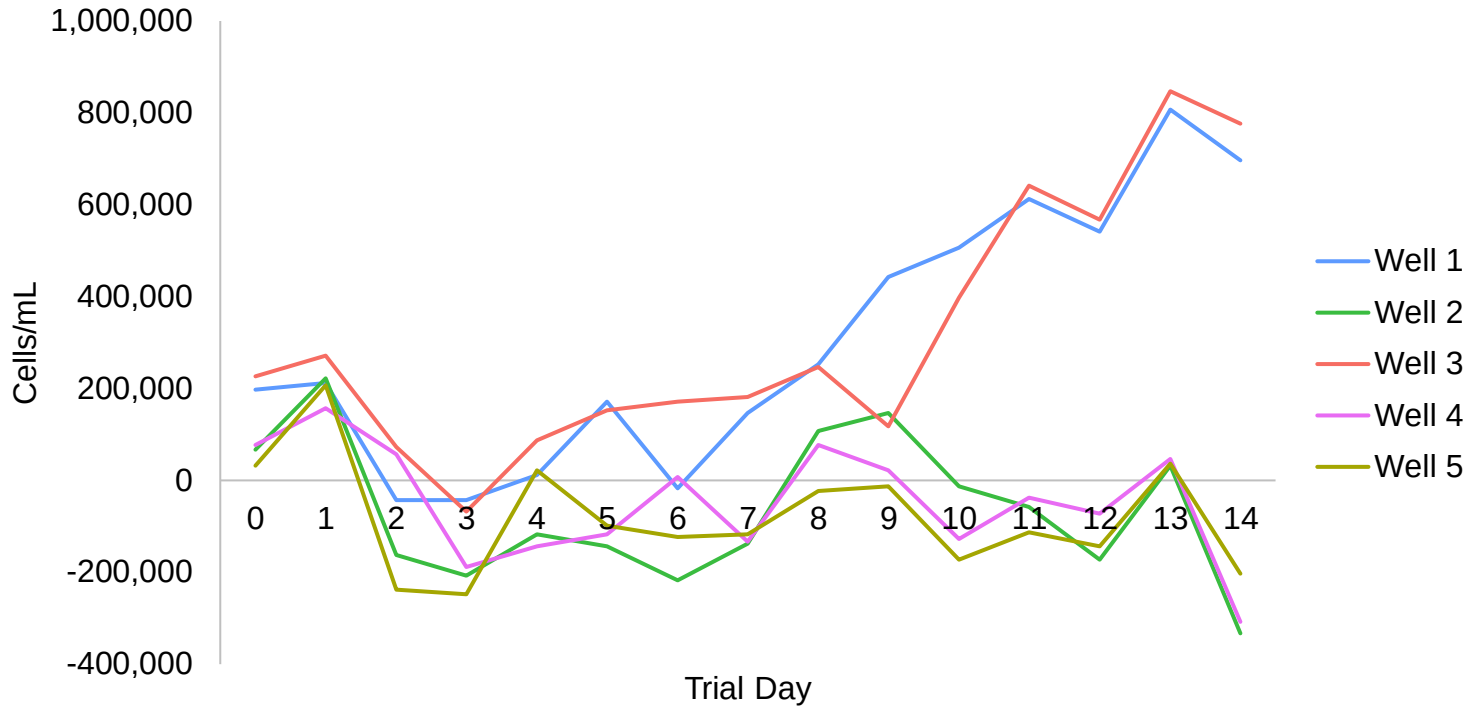

Figure S40 Growth responses of *Nitzschia frigida* culture in 96-well microplate under control conditions over 14 days. Measurements from five wells were taken daily using a microplate reader and converted into cells/mL based on standard curves.

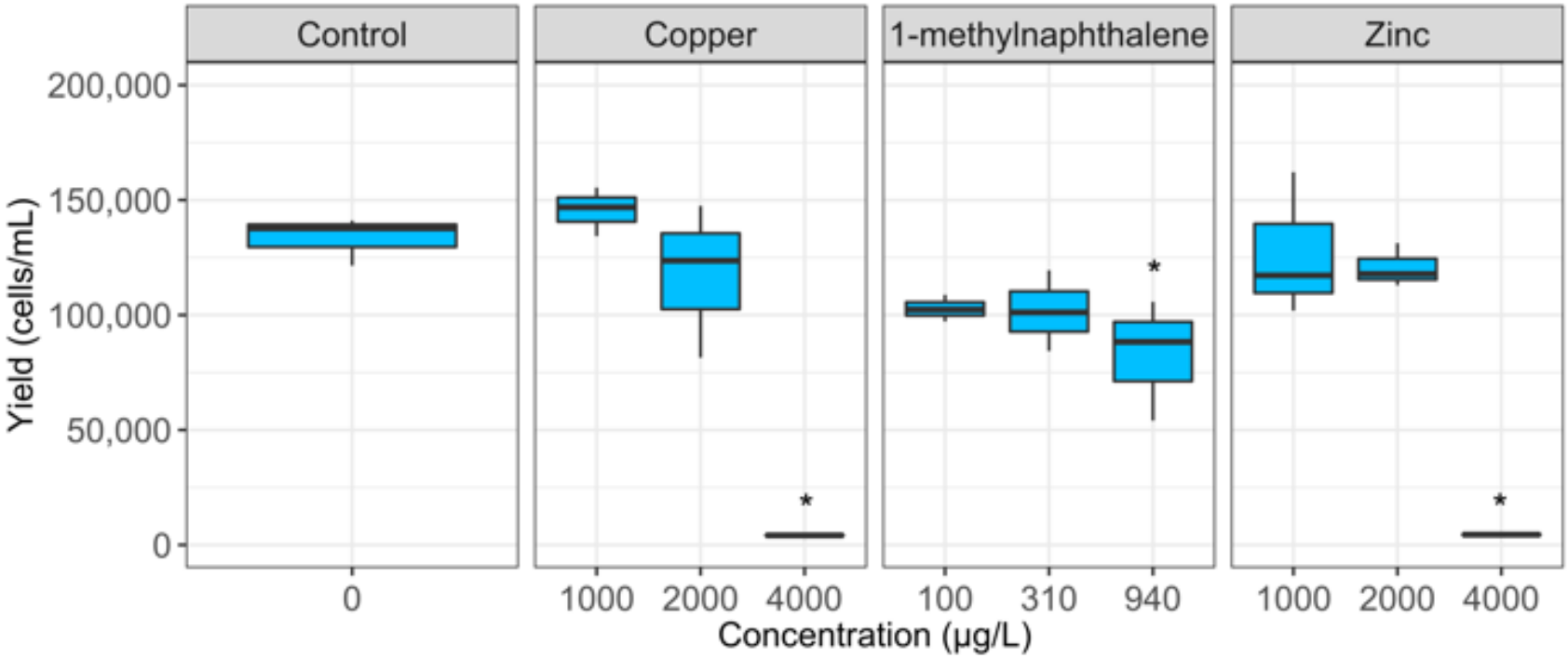

Figure S41 Boxplot of growth responses from 96-hour range-finding test conducted with *Skeletonema costatum*. Asterisks denote statistically significant differences ($p < 0.05$) from control based on one-way ANOVA and Dunnett's post-hoc test.

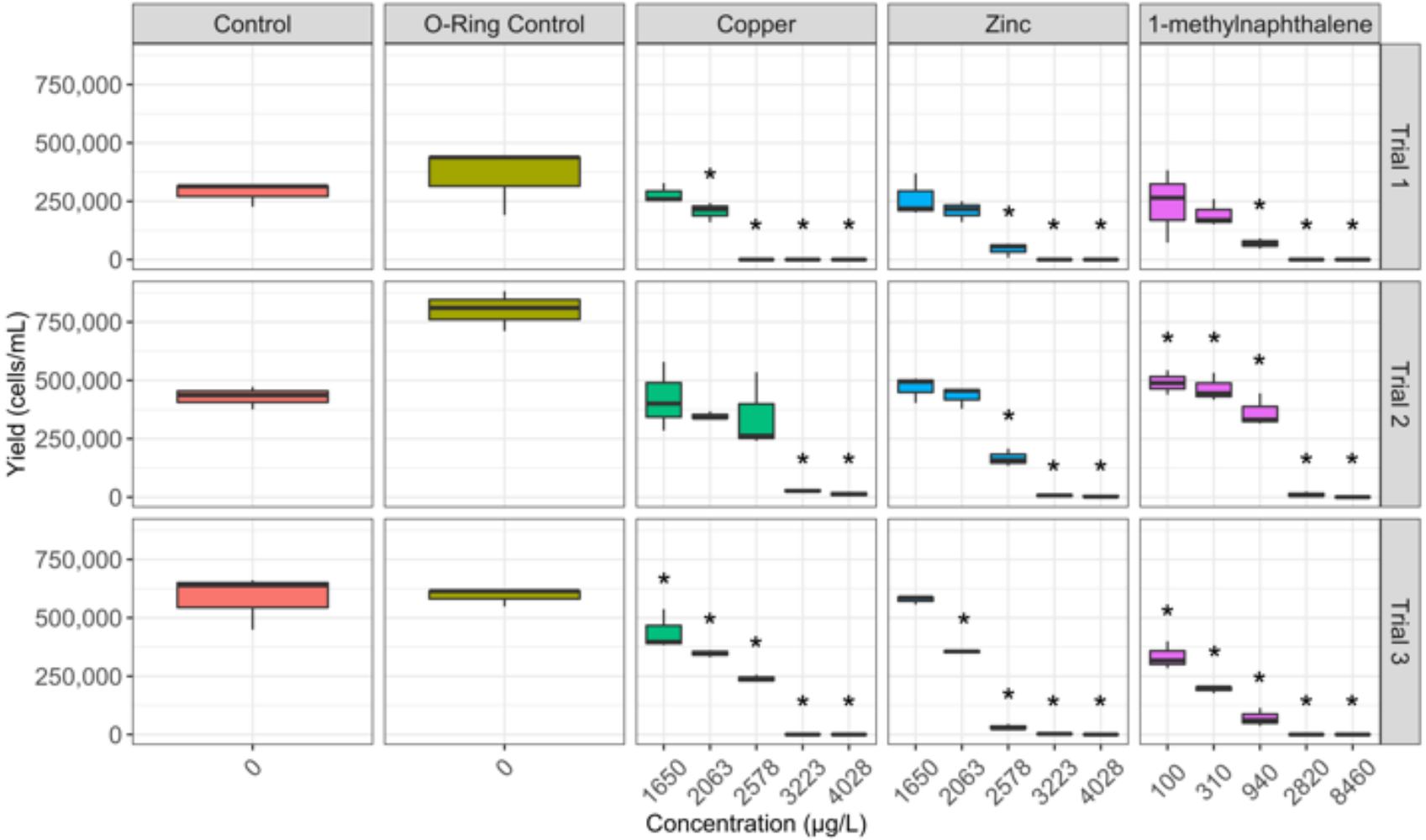

Figure S42 Box plots of 96-hour growth responses of *Skeletonema costatum* exposed to copper, zinc, and 1-methylnaphthalene. Asterisks denote statistically significant differences ($p < 0.05$) from controls based on one-way ANOVA followed by Dunnett's post-hoc test.

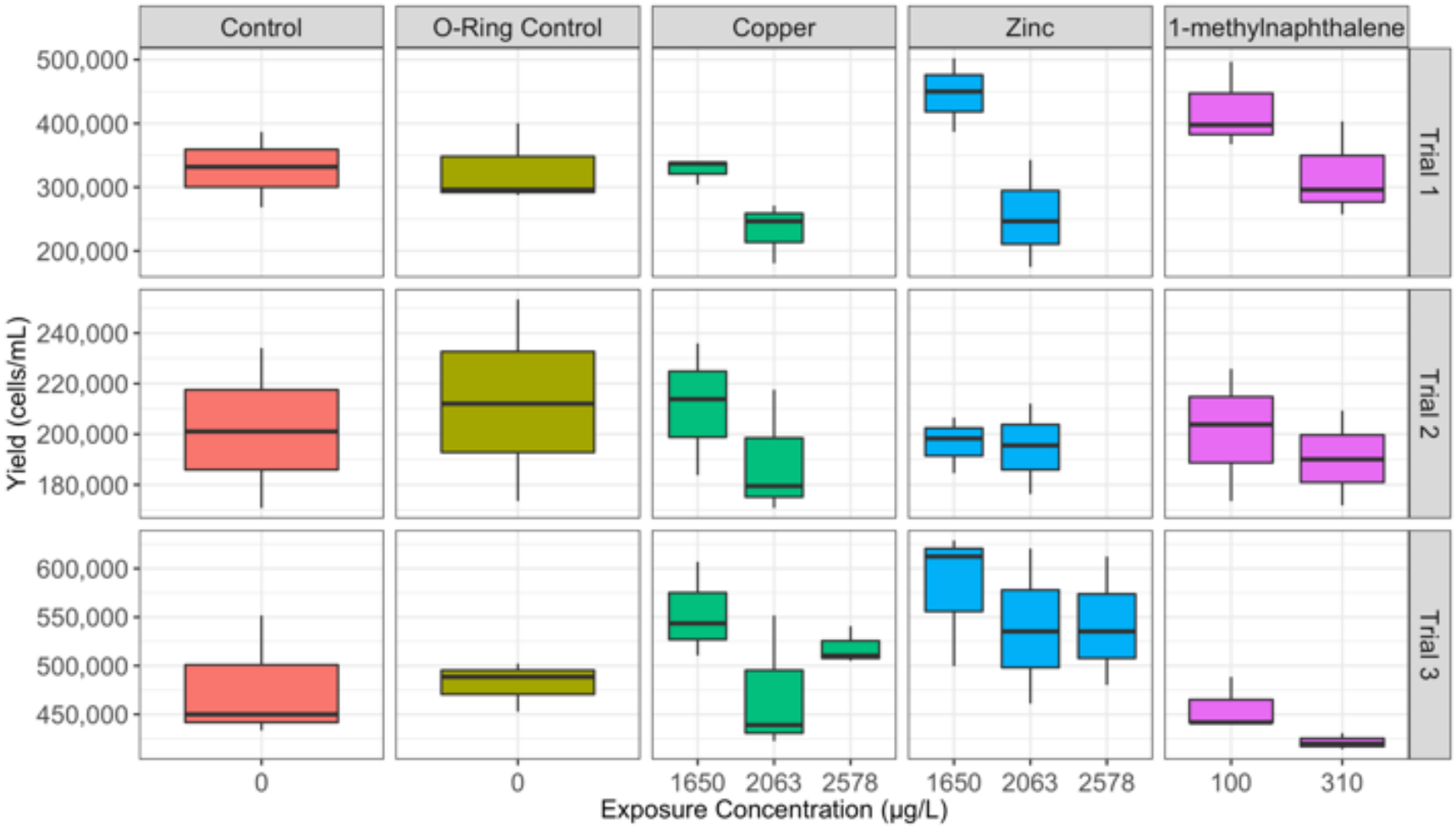

Figure S43 Box plots of growth responses of *Skeletonema costatum* after 96 days of exposure to various test substances followed by 96 days of recovery in clean media. No statistically significant differences ($p < 0.05$) from controls were observed based on one-way ANOVA followed by Dunnett's post-hoc test

5.3 References

Arctic Monitoring and Assessment Programme (AMAP). (2011). Mercury in the Arctic. *Arctic Monitoring and Assessment Programme, Oslo, Norway.*

Bejarano, A. (2018). Critical review and analysis of aquatic toxicity data on oil spill dispersants. *Environmental Toxicology and Chemistry, 37(*12), 2989-3001.

Bergmann, M., Mützel, S., Primpke, S., Tekman, M. B., Trachsel, J., Gerdts, G. (2019). White and wonderful? Microplastics prevail in snow from the Alps to the Arctic. *Science Advances, 5*(8), eaax1157.

Braune, B. M. (2007). Temporal trends of organochlorines and mercury in seabird eggs from the Canadian Arctic, 1975-2003. *Environmental Pollution, 148*(2), 599-613.

Buckman, A. H.; Veldhoen, N.; Ellis, G.; Ford, J. K. B.; Helbing, C. C.; Ross, P. S. (2011). PCB-associated changes in mRNA expression in killer whales (*Orcinus orca*) from the NE Pacific Ocean. *Environmental Science and Technology, 45*, 10194-10202.

Canadian Council of Ministers of the Environment (CCME). (1999). Canadian Water Quality Guidelines for the Protection of Aquatic Life: Chromium. Winnipeg: Canadian Council of Ministers of the Environment.

Canadian Council of Ministers of the Environment (CCME). (1999). Canadian Water Quality Guidelines for the Protection of Aquatic Life: Molybdenum. Winnipeg: Canadian Council of Ministers of the Environment.

Canadian Council of Ministers of the Environment (CCME). (2001). Canadian Water Quality Guidelines for the Protection of Aquatic Life: Arsenic. Winnipeg: Canadian Council of Ministers of the Environment.

Canadian Council of Ministers of the Environment (CCME). (2010). Canadian Water Quality Guidelines for the Protection of Aquatic Life: Ammonia. Winnipeg: Canadian Council of Ministers of the Environment.

Canadian Council of Ministers of the Environment (CCME). (2011). Canadian Water Quality Guidelines for the Protection of Aquatic Life: Chloride. Winnipeg: Canadian Council of Ministers of the Environment.

Canadian Council of Ministers of the Environment (CCME). (2014). Canadian Water Quality Guidelines for the Protection of Aquatic Life: Cadmium. Winnipeg: Canadian Council of Ministers of the Environment.

Canadian Council of Ministers of the Environment (CCME). (2018). Canadian Water Quality Guidelines for the Protection of Aquatic Life: Zinc. Winnipeg: Canadian Council of Ministers of the Environment.

Das, K., Dupont, A., De Pauw-Gillet, M.-C., Debier, C., Siebert, U. (2016). Absence of selenium protection against methylmercury toxicity in harbour seal leucocytes *in vitro. Marine Pollution Bulletin, 108*, 70-76.

Diercks, A.-R., Highsmith, R. C., Asper, V. L., Joung, D., Zhou, Z., Guo, L., Shiller, A. M., Joye, S. B., Teske, A. P., Guinasso, N., Wade, T. L., Lohrenz, S. E. (2010). Characterization of subsurface polycyclic aromatic hydrocarbons at the Deepwater Horizon site. *Geophysical Research Letters, 37*(20).

Dommergue, C. Larose., C. Fain, O. Clarisse, D. Foucher, H., Hintelmann, D. Schneider, C. P. Ferrari. (2010). Deposition of mercury species in the Ny-Ålesund area (79 degrees N) and their transfer during snowmelt. *Environmental Science and Technology, 44*, 901-907.

Durell, G., Utvik, T. R., Johnsen, S., Frost, T., Neff, J. (2006). Oil well produced water discharges to the North Sea. Part I: Comparison of deployed mussels (*Mytulis edulis*), semi-permeable membrane devices, and the DREAM model predictions to estimate the dispersion of polycyclic aromatic hydrocarbons. *Marine Environmental Research, 62*(3), 194-223.

Frouin, H., Loseto, L. L., Stern, G. A., Haulena, M., Ross, P. S. (2012). Mercury toxicity in beluga whale lymphocytes: Limited effects of selenium protection. *Aquatic Toxicology, 109,* 185-193.

Grenvald, J. C., Nielsen, T. G., Hjorth, M. (2013). Effects of pyrene exposure and temperature on early development of two co-existing Arctic copepods. *Ecotoxicology, 22*, 184-198.

Hansen B.H., Altin D., Rørvik S. Øverjordet, I., Jager T., and Nordtug T. (2013). Acute exposure of water soluble fractions of marine diesel on Arctic *Calanus glacialis*

and boreal *Calanus finmarchicus*: Effects on survival and biomarker response. *Science of the Total Environment, 449,* 276-284.

Mares – Scuba Diving Blog. (2017, March 30). *Melosira arctica* named algae species of 2016. Mares. Retrieved November 17, 2021 from https://blog.mares.com/melosira-arctica-named-algae-species-of-2016-1153.html.

Moore, D., Sibley, P. K., Solomon, K. R. (2016). *Toxicity of salts, metals, and nitrogenous contaminants to cold-water fish under northern conditions*. (Unpublished doctoral dissertation). University of Manitoba, Winnipeg, MB.

Negri, A., Luter, H., Fisher, R., Brinkman, D., Irving, P. (2018). Comparative toxicity of five dispersants to coral larvae. *Scientific Reports, 8*, 3043.

Sammarco, P. W., Kolian, S. R., Warby, R. A. F., Bouldin, J. L., Subra, W. A., Porter, S. A. (2013). Distribution and concentrations of petroleum hydrocarbons associated with the BP/Deepwater Horizon oil spill, Gulf of Mexico. *Marine Pollution Bulletin, 73*(1), 129-143.

Schulz-Baldes, M. (1974). Lead uptake from sea water and food, and lead loss in the common mussel *Mytulis edulis*. *Marine Biology, 25*, 177-193.

Szczybelski, A. S., van den Heuvel-Grave, M. J., Koelmans, A. A., van den Brink, N. W. (2019). Biomarker responses and biotransformation capacity in Arctic and temperate benthic species exposed to polycyclic aromatic hydrocarbons. *Science of the Total Environment, 662,* 631-638.

Wade, T. L., Sericano, J. L., Sweet, S. T., Knap, A. H., Guinasso, N. L. (2016). Spatial and temporal distribution of water column total polycyclic aromatic hydrocarbons (PAH) and total petroleum hydrocarbons (TPH) from the Deepwater Horizon (Macondo) incident. *Marine Pollution Bulletin, 103*(1-2), 286-293.

Wells, P. (1984). The toxicity of oil spill dispersants to marine organisms: a current perspective. Oil Spill Chemical Dispersants: Research, Experience, and Recommendations. *ASTM International,* 177-202.