Received: 20 February 2018 | Revised: 20 April 2018 | Accepted: 23 April 2018

DOI: 10.1002/ece3.4213

ORIGINAL RESEARCH

WILEY Ecology and Evolution Open Access

eDNA metabarcoding as a new surveillance approach for coastal Arctic biodiversity

Anaïs Lacoursière-Roussel[1] | Kimberly Howland[2] | Eric Normandeau[3] | Erin K. Grey[4] | Philippe Archambault[5] | Kristy Deiner[6] | David M. Lodge[7] | Cecilia Hernandez[3] | Noémie Leduc[3] | Louis Bernatchez[3]

[1]St. Andrews Biological Station (SABS), Fisheries and Oceans Canada, St. Andrews, NB, Canada

[2]Central and Arctic Region, Fisheries and Oceans Canada, Freshwater Institute, Winnipeg, MB, Canada

[3]Department of Biology, Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec, QC, Canada

[4]Division of Science, Mathematics and Technology, Governors State University, University Park, IL, USA

[5]Department of Biology, Université Laval, Québec, QC, Canada

[6]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zürich, Switzerland

[7]Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY, USA

Correspondence
Anaïs Lacoursière-Roussel, St. Andrews Biological Station (SABS), Fisheries and Oceans Canada, St. Andrews, NB, Canada.
Email: anais.lacoursiere.1@ulaval.ca

Funding information
ArcticNet; POLAR knowledge; Nunavut Wildlife Management Board; Fisheries and Oceans Canada Aquatic Invasive Species Monitoring Program

Abstract

Because significant global changes are currently underway in the Arctic, creating a large-scale standardized database for Arctic marine biodiversity is particularly pressing. This study evaluates the potential of aquatic environmental DNA (eDNA) metabarcoding to detect Arctic coastal biodiversity changes and characterizes the local spatio-temporal distribution of eDNA in two locations. We extracted and amplified eDNA using two COI primer pairs from ~80 water samples that were collected across two Canadian Arctic ports, Churchill and Iqaluit based on optimized sampling and preservation methods for remote regions surveys. Results demonstrate that aquatic eDNA surveys have the potential to document large-scale Arctic biodiversity change by providing a rapid overview of coastal metazoan biodiversity, detecting nonindigenous species, and allowing sampling in both open water and under the ice cover by local northern-based communities. We show that DNA sequences of ~50% of known Canadian Arctic species and potential invaders are currently present in public databases. A similar proportion of operational taxonomic units was identified at the species level with eDNA metabarcoding, for a total of 181 species identified at both sites. Despite the cold and well-mixed coastal environment, species composition was vertically heterogeneous, in part due to river inflow in the estuarine ecosystem, and differed between the water column and tide pools. Thus, COI-based eDNA metabarcoding may quickly improve large-scale Arctic biomonitoring using eDNA, but we caution that aquatic eDNA sampling needs to be standardized over space and time to accurately evaluate community structure changes.

KEYWORDS

arctic, coastal biodiversity, eDNA metabarcoding, global changes, invasion, spatio-temporal distribution

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2018 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

In the Arctic, climate change and marine invasions are expected to result in over 60% species turnover from present biodiversity with substantial impacts on marine ecosystems (Cheung et al., 2009). Climate change is opening new waterways in the Arctic Ocean, resulting in greater shipping traffic (ACIA 2004; Arctic Council 2009; Guy & Lasserre, 2016). Predicted increases in shipping frequency and routes (Eguíluz, Fernández-Gracia, Irigoien, & Duarte, 2016; Miller & Ruiz, 2014; Smith & Stephenson, 2013), increased infrastructure development in ports (Gavrilchuk & Lesage, 2014), and associated chemical/biological pollution will place other ecosystem services at risk. Furthermore, the introduction of nonindigenous species (NIS) may displace native species, alter habitat and community structure and increase aquaculture and fishing gear fouling in estuaries and coastal zones (Goldsmit et al., 2018; Grosholz, 2002; Parker et al., 1999). Currently, the continuous monitoring needed to evaluate large-scale changes in coastal biodiversity and faunal assemblages in the Canadian Arctic is limited (Archambault et al., 2010), hindering risk management and ecosystem sustainability planning (Larigauderie et al., 2012).

Recent advances in the collection and analysis of environmental DNA (eDNA) provide a new complementary approach that can help to fill gaps in regional species distribution data left by logistically difficult traditional methods (e.g., bottom trawl, SCUBA diving) (Deiner et al., 2017), particularly in remote and otherwise challenging locations. eDNA allows for the detection of traces of DNA in water from macro-organisms (Thomsen, Kielgast, Iversen, Wiuf, et al., 2012). Collecting water samples for eDNA surveys could allow rapid sample collection, reduce the cost associated with data collection/shipping, and is less destructive because it does not require the manipulation of organisms (Lodge et al., 2012; Taberlet, Coissac, Hajibabaei, & Rieseberg, 2012). eDNA metabarcoding (i.e., high-throughput eDNA sequencing) can enable the identification of millions of DNA fragments/sample, providing a powerful approach to survey aquatic biodiversity. Repeated eDNA surveys could potentially be used to evaluate long-term biodiversity changes such as detecting native species loss and declines, NIS introductions and range expansions, and community structure changes. However, the detection of species using eDNA varies as a function of the population densities (Lacoursière-Roussel, Côté, Leclerc, & Bernatchez, 2016; Lacoursière-Roussel, Dubois, & Bernatchez, 2016; Mahon et al., 2013), life history traits, shedding rates (Lacoursière-Roussel, Rosabal, & Bernatchez, 2016; Sassoubre, Yamahara, Gardner, Block, & Boehm, 2016) local environmental conditions and technical approaches such as sequencing efforts and primer biases (Freeland, 2017; Pawluczyk et al., 2015). Moreover, major concerns with eDNA metabarcoding, including its ability to accurately identify sequences to species (Chain, Brown, MacIsaac, & Cristescu, 2016) and the unknown ecological dynamics of eDNA in coastal ecosystems, need to be studied before marine biodiversity can be compared across spatial and temporal scales using this method.

Little is currently known about the efficacy of eDNA metabarcoding in surveying long-term variation in marine coastal biodiversity (Lim et al., 2016; Port et al., 2016; Thomsen & Willerslev, 2015). Relative to freshwater ecosystems where more studies have been conducted, eDNA in coastal marine ecosystems is diluted into a much larger volume of water and exposed to pronounced hydrodynamics (e.g., tides, currents) and variation in abiotic conditions (e.g., salinity, temperature), which is likely to affect eDNA transport and degradation (Foote et al., 2012; Thomsen, Kielgast, Iversen, Møller, et al., 2012). In spite of these challenges, a recent study of horizontal spatial eDNA distribution in the Puget Sound (Washington, USA; O'Donnell et al., 2017) was successful in revealing fine scale distribution of species in these communities. In Arctic ecosystems, higher eDNA transport and diffusion is expected due to slower DNA degradation in cold-water temperatures, but no study has yet characterized aquatic eDNA distribution in this environment. Improving our understanding of the ecology of eDNA—the myriad of interactions between extraorganismal genetic material and its environment (Barnes & Turner, 2016)—in various ecosystems is fundamental to determining how eDNA can and cannot improve biodiversity research.

Our objective is to explore the potential of eDNA as a biodiversity monitoring approach to assist in rapid detection of coastal biodiversity shifts on large spatial scale in two Arctic coastal areas: Churchill and Iqaluit. These two Arctic commercial ports are expected to be particularly prone to biodiversity changes because they are among the top three ports in the Canadian Arctic with respect to vessel arrivals and associated ballast and/or hull fouling invasions risk (Chan, Bailey, Wiley, & MacIsaac, 2013). More specifically, we estimate the proportion of the Arctic biodiversity that can be identified at the species level with eDNA, and we then characterize the spatio-temporal distribution of eDNA with respect to water column depths, tide pools, and seasons.

2 | MATERIALS AND METHODS

The spatio-temporal eDNA distribution was characterized at three different depths in the water column, in tide pools, and between summer and fall seasons. Specifically, water samples were collected in 13 subtidal sites at three different depths (surface, middepth and deep water (i.e., 50 cm from the bottom), 12 tide pool sites within three intertidal areas (N = 4 sites/area) and 20 samples were collected at a single site from the shore approximately 2 m spaced along a transect (Figure 1). For the summer period (without ice cover), Churchill and Iqaluit were surveyed in 2015 between August 11–14 and August 17–22, respectively (hereafter called S20). To evaluate seasonal effects (Iqaluit only), the 20 samples at a single site were collected during fall (November 18th, 2015) near shore from water that rose between ice pans at high tide (hereafter called F20).

Each sample (250 ml water) was collected using a Niskin bottle and then rapidly filtered in the field through a 0.7 μm glass microfiber filter (Whatman GF/F, 25 mm) using syringes (BD 60 ml; Kranklin Lakes, NJ, USA). Field negative controls (i.e., 250 ml distilled water)

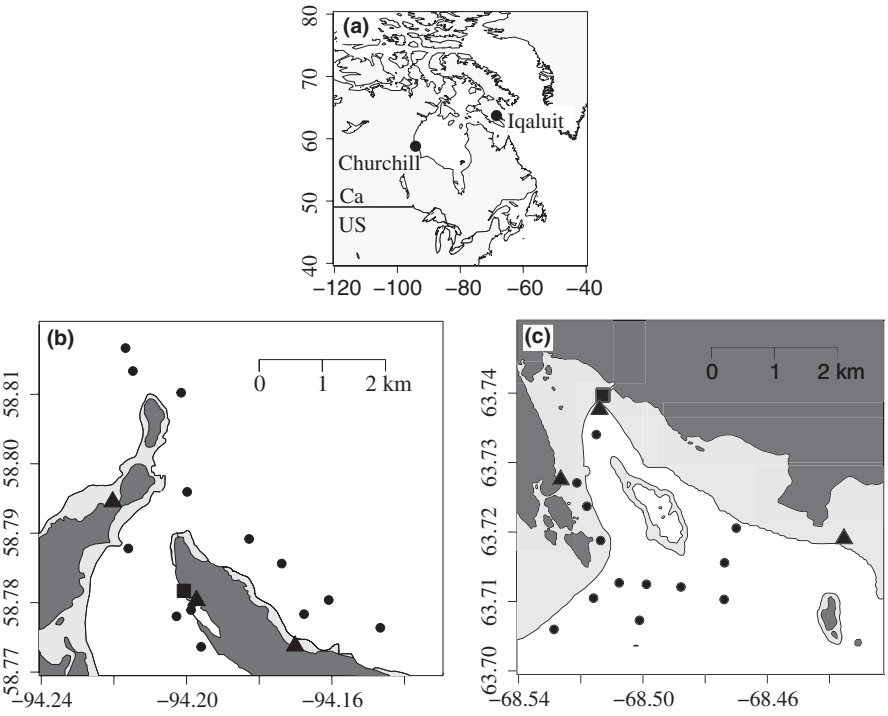


FIGURE 1 Geographical locations of the sampling port in the Canadian Arctic (map a) and the site distribution within Churchill (map b) and Iqaluit (map c). Subtidal areas are shown in white and the intertidal areas in light gray. Circles depict the water column sites, triangles are the tide pools sites and the squares are the S20 and F20 shore sampling sites

were filtered for every 10 samples. Filters were preserved at 4°C in 700 µl of Longmire's lysis/preservation buffer within a 2 ml tube for up to 3 weeks (Wegleitner, Jerde, Tucker, Chadderton, & Mahon, 2015) and then frozen at −20°C until DNA extraction. To reduce risk of crosscontamination during sampling and the filtration process, individual sampling kits were used for each sample (bottles and filter housing sterilized with a 10% bleach solution and new sterilized gloves, syringes, and tweezers). Each sampling kit was exposed to UV for 30 min. To reduce the risk of laboratory crosscontamination, procedures for eDNA extraction, PCR preparation, and post-PCR steps were all performed in different rooms and PCR manipulations were performed in a decontaminated UV hood. Samples from a specific port were all treated together, and the bench space and laboratory tools were bleached and exposed to UV for 30 min prior to processing the next port. Sites within a port were processed in a randomized order.

2.1 | eDNA extraction, amplification and sequencing

DNA was extracted using a QIAshredder and phenol/chloroform protocol (see Supporting Information Appendix S1). Negative control extractions (950 µl distilled water) were performed for each sample batch (i.e., one for each 23 samples). Two pairs of universal metazoan mitochondrial cytochrome c oxidase subunit I (COI) primers that have been developed and tested on a broad array of marine species were used to amplify eDNA from as many metazoan taxa as possible: the forward mlCOIintF (Leray et al., 2013) and reverse jgHCO2198 (Geller, Meyer, Parker, & Hawk, 2013) amplifying 313 bp (hereafter called COI1) and the forward LCO1490 (Folmer, Black, Hoeh, Lutz, & Vrijenhoek, 1994) and reverse ill_C_R (Shokralla et al., 2015) amplifying 325 bp (COI2).

The performance of the two selected primer pairs used in this study was previously tested on 104 zooplankton species and was validated on mock metazoan communities collected in Canadian ports by Zhang (2017). Based on a total of 13 COI primer pairs selected from the literature and tested, Zhang (2017) showed the efficiency of using multiple COI primer pairs in a single Illumina run to recover species by metabarcoding and detected 32% of species using COI1 and 49% of species using COI2. Here, the DNA amplification protocols for both primer pairs were optimized in vitro using 12 Arctic specimens and 12 potential invaders (i.e., annealing temperature gradient using DNA extracted from tissue samples; Supporting Information Table S1). The primer sequences and sequence databases were also evaluated in silico for their ability to detect native and potential nonindigenous Arctic metazoans. A list of recorded coastal Arctic metazoans was obtained by pooling all Arctic species databases that we had access to (*N* total = 897 metazoan identified at the species level; Fisheries and Oceans Canada Arctic Marine Invertebrate Database (Supporting Information Appendix S2), Archambault unpublished data, Cusson, Archambault, and Aitken (2007), Goldsmit, 2016; Goldsmit, Howland, & Archambault, 2014; K. Howland, P. Archambault, N. Simard and R Young, unpublished data, Piepenburg et al., 2011; Link, Piepenburg, & Archambault, 2013; López, Olivier, Grant, & Archambault, 2016; Olivier, San Martín, & Archambault, 2013; Roy, Iken, & Archambault, 2015; Young, Abbott, Therriault, & Adamowicz, 2016). Potential NIS invaders (*N* = 130 species) were targeted based on (1) screening level risk assessments and predictive species distribution models indicating they were high risk (Goldsmit et al., 2017), (2) their presence in ports connected to the Canadian Arctic, and/or (3) their presence in ballast waters and hulls of ships based on monitoring at Canadian Arctic ports (Chan, MacIsaac, & Bailey, 2015; Chan et al., 2012). Historical data include many Arctic

regions, surveyed mainly during the open water period, with focal taxa varying among surveys. Comprehensive port surveys in Churchill and Iqaluit were only conducted once every few years (Churchill 2007, 2011 and 2015; Iqaluit 2012 and 2015–2016). A script was used to determine whether the primer sequences were present for the targeted species (species previously recorded from the Artic and potential NIS) available in the NCBI and BOLD databases (September 2016; http://www.barcodinglife.org). Searches for Arctic species in the sequence databases were performed with Python and Bash programs (developed by Jérôme Laroche at the Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval) and analyses are freely available on Bitbucket (https://bitbucket.org/jerlar73/env-dna).

Three PCR replicates were performed for each eDNA sample and each primer set. DNA amplifications were performed in a one-step dual-indexed PCR approach designed for Illumina instruments at IBIS. The final reaction volume for each PCR replicate was 24 μl; including 12.5. μl Qiagen Multiplex Mastermix, 6.5 μl diH_20, 1 μl of each primer (10 μM), and 3.0 μl of DNA. For all samples, the PCR mixture was denatured at 95°C for 15 min, followed by 35 cycles (94°C for 30 s, 54°C for 90 s (except for the COI2 primers, which were at 52°C for 90 s and 72°C for 60 s) and a final elongation at 72°C for 10 min. Products of the three aliquots were pooled for each sample. A negative PCR control was performed for each sample and primer set. All amplifications were visualized on a 1.5% agarose gel electrophoresis. No positive amplification of the PCR negative control was observed. Field and extraction negative controls were treated exactly the same as regular samples and were also sequenced. Pooled products were purified using Axygen PCR clean up kit following the manufacturer's recommended protocol. Libraries were quantified by AccuClear Ultra High Sensitivity dsDNA Quantification Kit using the TECAN Spark 10 M Reader for each sample and were pooled in equal molar concentrations to maximize equal sequence depth per sample location (150 and 37 ng per sample for COI1 and COI2 primer sets, respectively, in Churchill and 200 and 300 ng per sample for COI1 and COI2 primer sets, respectively, in Iqaluit). When Quant-iT PicoGreen (Life Technologies) did not detect any DNA, 22.0 μl PCR mixtures were mixed nonetheless (see Supporting Information Table S2 for the concentration and volume for each sample separately).

Sequencing was carried out using an Illumina MiSeq (Illumina, San Diego, CA, USA) using a paired-end MiSeq Reagent Kit V3 (Illumina) and following the manufacturer's instructions (Supporting Information Appendix S1). Each port was analyzed on a separate run to ensure independency, but the samples within a port were pooled within a single Illumina MiSeq run to ensure the equality of sequencing depth among samples. Raw sequences reads were deposited in NCBI's Sequence Read Archive (SRA, http://www.ncbi.nlm.nih.gov/sra) under Bioproject PRJNA388333.

2.2 | Taxonomic identification

Forward and reverse sequences for each sample were trimmed using Trimmomatic 0.30 (Bolger, Lohse, & Usadel, 2014). FastQC version v0.11.3 was used to confirm the quality of the trimmed reads (Andrews, 2010). The Fastq quality scores were all well above 20 for the trimmed reads. Reads were then merged using FLASH v1.2.11 with a minimum overlap of 30 bp (Magoč & Salzberg, 2011). "Orphan" reads with <30 bp of overlap between forward and reverse reads were discarded and only merged reads were used in the analyses. COI1 and COI2 amplicons were split using a Python script which searches for degenerate primers at the beginning and end of each sequence and only keeps sequences where there is positive identification for both primers ≥270 bp. These sequences were compared for identity with the metazoan sequences present in the Barcode of Life Database (BOLD) (Ratnasingham & Hebert, 2007; available on the BOLDSYSTEM S3 website, http://www.boldsystems.org, on the 22nd August 2016). Terrestrial species (insects, human, birds, and mammals) and sequences that did not have a taxonomic name assigned at the species level were removed from the reference database.

To examine biodiversity at the species level, direct taxonomic assignment of each merged read with ≥97% identity was performed using the *Barque* pipeline version 0.9 (see Supporting Information Appendix S3), an open source and freely available metabarcoding analysis pipeline (www.github.com/enormandeau/barque). Reads matching with equal quality scores to more than one species due to low interspecific divergence were found using usearch. Only 156 reads (i.e., 0.02% reads, 17 cases) in total were found with such multiple hits. For each case, the list of species was scrutinized and species that were clearly not expected in the Arctic based on Ocean Biogeographic Information system (OBIS), The World Porifera Database, the World Register of Marine Species (WoRMS) database, invasion risk assessments (see references above and Supporting Information Appendix S2), and expert knowledge were removed from the sequence reference database mentioned above (see Supporting Information Table S3 for details about the multiple hits and actions made for each species). The pipeline was run again to find the top hits only. The proportion of missing species assignments due to BOLD incompleteness was further explored for each metazoan phyla using Operational Taxonomic Units (OTU) clustering according to 97% similarity with swarm 2.2.0 (Mahé, Rognes, Quince, De Vargas, & Dunthorn, 2015; see bioinformatic details Supporting Information Appendix S3). OTUs represented by a single read (singletons) were excluded and the identity between the representative sequences and the BOLD database was performed using vsearch (Rognes, Flouri, Nichols, Quince, & Mahé, 2016). For each phylum, proportion of the biodiversity assigned to the species level was obtained from the number of OTUs between 97–100% (similar to threshold used to assign species for sequences in the BOLD database) relative to those between 80–97% (i.e., below species level).

2.3 | Statistical analyses

Sampling effort is an important factor to consider in both traditional and eDNA biodiversity surveys. Two levels of port-specific sampling effort were explored: number of unique species per read (a measure of sequencing effort) and the number of unique species per sample

(a measure of eDNA collection effort). For water column (surface, middepth and deep), tide pool and shore (S20 and F20) sampling locations, we plotted both read and sample rarefied accumulation curves to visualize whether or when a plateau was reached (which would indicate adequate sequencing and sampling effort to characterize all species). We also inspected the relative position of the read curve compared to the sample curve, as read curves lying above sample curves typically indicate spatial aggregation of species (Gotelli & Colwell, 2010), or in this case eDNA sequences. These sampling effort analyses were performed in R 3.4.1 using the *specaccum* function in the *vegan* package.

All further statistical analyses were performed using R 3.0.3. The spatial distribution of eDNA and the seasonal variability in the community composition was represented using Principal component analysis (PCoA) and tested using PERMANOVA (Anderson, 2001) after Hellinger transformation. Hellinger transformation was appropriate to deal with the large proportion of zeros and reduces the importance of large abundances (Legendre & Legendre, 1998) that could be due to the eDNA origin (e.g., capture of cell or mitochondria vs. extracellular DNA) or the amplification process. Species that mostly contributed to the dissimilarity/similarity between the treatments (depths and tide pools vs. water column) were identified using SIMPER analysis using the *simper()* function of the *vegan* package. Shannon diversity indices were calculated with the R package vegan. Analyses of variance (ANOVAs) were used to test whether species diversity, richness and $\log_{10}$(reads abundance) varied as a function of sampling location (i.e., water column and tide pools; sites included as a random variable) and water depths for each port separately using the *lme()* function of the *NLME* package (Pinheiro, Bates, DebRoy, & Sarkar, 2017) with sites included as a random variable (interactions between sites and depths could not be tested due to unique values per depth). The seasonal effect on read abundance (i.e., metazoan reads, see section *taxonomic identification*), Shannon diversity and species richness was evaluated using a Student's *t* test comparing the S20 and F20 samples in Iqaluit. Sørensen and Jaccard nonparametric estimates were calculated for location pairs using the *SimilarityPair* function of the *SpadeR* package in R (Chao, Ma, Hsieh, & Chiu, 2016) to test for the level of similarity in species composition between sampling location and seasons.

3 | RESULTS

After bioinformatics filtering (Supporting Information Table S2), we obtained 712,494 aquatic eukaryotic reads in Churchill (200,732 reads for COI1 and 511,762 reads for COI2) and 178,728 reads in Iqaluit (100,139 reads for COI1 and 78,589 reads for COI2). No amplification was visualized on the gel electrophoresis for the negative PCR controls and no significant eDNA reads were sequenced in any of the negative extractions controls (Churchill: 1–12 reads, average of 0.05% of the eDNA sample reads; Iqaluit: 1–8 reads, average of 0.17% of the eDNA samples reads) or the negative field controls (Churchill: 2–73 reads, 0.30% in average of the eDNA sample reads; Iqaluit: 0–54 reads, 0.75% in average of the eDNA sample reads).

Cytochrome c oxidase subunit I sequences of 46% and 44% of the known Canadian Arctic native taxa and 63% and 53% of

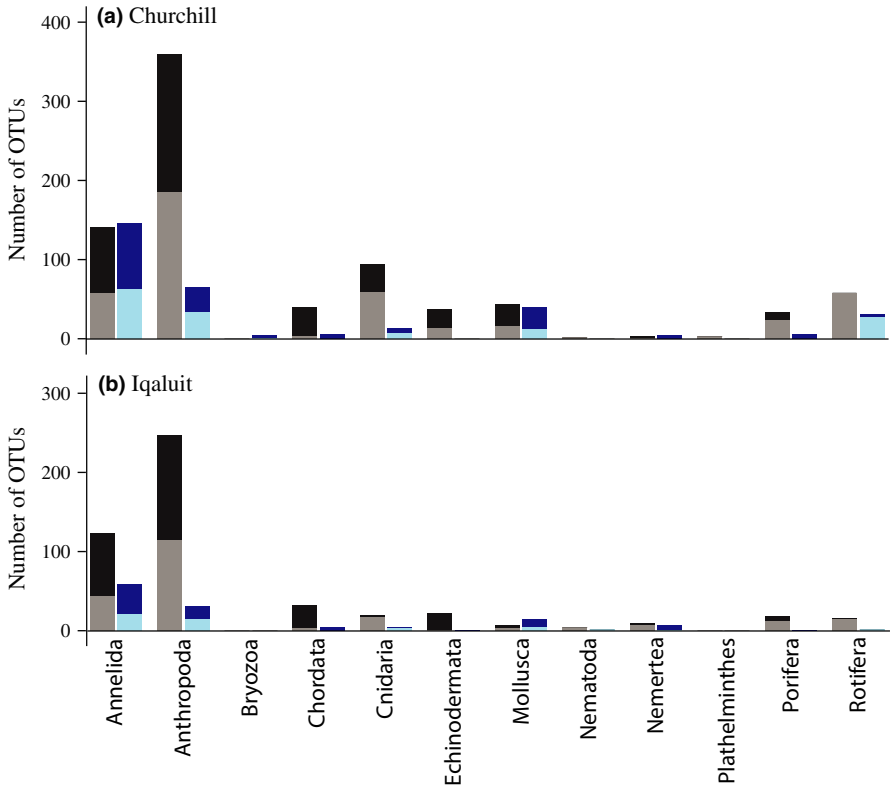


FIGURE 2 The number of Operational taxonomic units (OTU) identified at the species level (dark: ≥97% identity) relative to those identified below the species level (lighten: ≥85% and <97% identity) for each phylum and from the COI1 (mlCOIintF-jgHCO2198: black and gray) and COI2 (LCO1490-ill_C_R: blue) primer sets separately for both Arctic sampling ports (Churchill and Iqaluit)

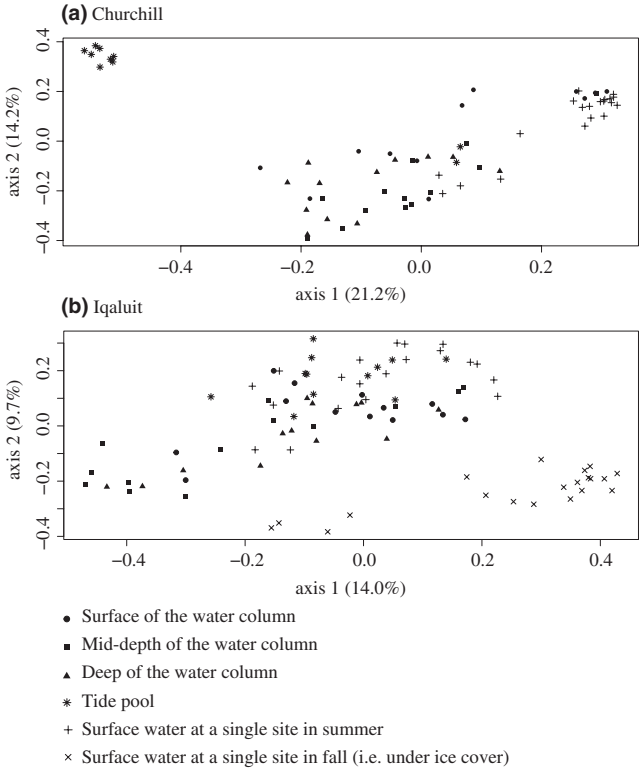


FIGURE 3 Principal component analysis depicting the community structure at the species level among sampling locations: water column (surface, middepth and deep water), tide pools (i.e., intertidal zone) and surface water collected in a single site in summer (i.e., S20) and in fall (F20) for both Arctic sampling ports (Churchill and Iqaluit). Ports were analyzed separately because each port was treated on a separate sequencing run

potential invaders are currently in GenBank or BOLD database, respectively. In parallel, the proportion of OTUs matched to a species in the eDNA survey was 53% in Churchill and 50% in Iqaluit (see the proportion by phylum in Figure 2). For both ports, the sampling effort could have been increased to reveal additional species as the sample and read accumulation curves did not plateau (Supporting Information Figure S1). However, there was little evidence for spatial eDNA aggregation within a location as sample-based curves fell only slightly below read curves, and within 95% confidence intervals, at all locations.

3.1 | Taxonomic composition in Arctic coastal ports

A total of 181 species were detected in the eDNA survey; 140 species in Churchill and 87 species in Iqaluit (see Supporting Information Figure S2 for the species list for each primer set and their status according to previous Canadian Arctic reports). Forty-eight species were amplified with both COI primer sets, 116 species recorded by the COI1 primer set only and 17 species by the COI2 primer set. At the species level, the primer sets detected a total of ten phyla; including nine phyla for the COI1 primer set (44 Annelida species, 31 Arthropoda, 35 Chordata, 17 Cnidaria, 17 Echinodermata, eight Mollusca, three Nemertea, five Porifera and four Rotifera) and 10 for the COI2 primer set (27 Annelida species, ten Arthropoda, two Bryozoa, five Chordata, six Cnidaria, one Echinodermata, eight Mollusca, two Nemertea, three Porifera and one Rotifera). In contrast to mock metazoan communities (see method section), a larger number of species was identified using COI1 primers than COI2 primers, but the latter detected proportionately more Annelida and Porifera.

For both ports, 74.0% of the species detected have been previously reported from the Arctic (Churchill: 70.0% and Iqaluit: 87.4%; COI1: 78.6% and COI2: 61.5%). The number of species detected using eDNA in Churchill and Iqaluit represents 10.9% and 8.5% metazoan species recorded within the overall Arctic species databases. Forty-seven species not previously reported were detected, including 15 Annelida, five Arthropoda, two Bryozoa, four Chordata, eight Cnidaria, two Echinodermata, four Mollusca, three Nemertea and four Rotifera species. The only potential invaders detected, the Arthropoda *Acartia tonsa*, was found with the COI1 primers in Churchill (64 reads averaging 99.4% identity with the sequence references). This species was previously recorded in ballast water in ports connected to Churchill and is considered a potential invader (Chan et al., 2012). However, COI sequences in BOLD assigned to *A. tonsa* are not monophyletic and several are indistinguishable from sequences assigned to the native *A. hudsonica*, suggesting misidentification of some *Acartia* specimens in BOLD.

3.2 | Spatial eDNA distribution

For both ports, the community structure differed significantly between the water column and the tide pools, but the proportion of explained variance was greater for Churchill than Iqaluit (Figure 3, PERMANOVA; Churchill: $R^2 = 0.21$, $p < 0.001$; Iqaluit: $R^2 = 0.12$, $p < 0.001$; seasonality did not impact analysis of spatial variability when analyzed separately). For both ports, the water column was dominated by Arthropoda (Churchill: 91,219 reads for COI1 and 164,080 reads for COI2; Iqaluit: 30,550 reads for COI1 and 16,971 reads for COI2), followed by Annelida (Churchill: 28,607 reads for COI1 and 110,643 reads for COI2; Iqaluit: 11,518 reads for COI1 and 2,621 reads for COI2) (Figure 4). Mollusca species were mainly detected in tide pools at both ports (91% and 23%, respectively, for Churchill and Iqaluit; Figure 4), and were by far the dominant taxa in Churchill with the majority being *Littorina saxatilis* for COI1 and COI2 (95.8% (i.e., 14,219 reads) and 100% (i.e., 198,684 reads) of Mollusca reads; cumulative contributions for Churchill = 62.4% and Iqaluit = 52.0%); tide pools were dominated by Arthropoda species in Iqaluit (Figure 4).

The Shannon diversity index was significantly greater in the water column than tide pools in Churchill (ANOVA: $p = 0.002$), but there was no significant difference in Iqaluit ($p = 0.2$; Figure 5). In Churchill, despite a significantly greater number of reads in tide pools than the water column (averaging 23,276 and 11,623 reads in tide pools and water column samples, respectively; $p = 0.06$), there was no significant difference in species richness between water

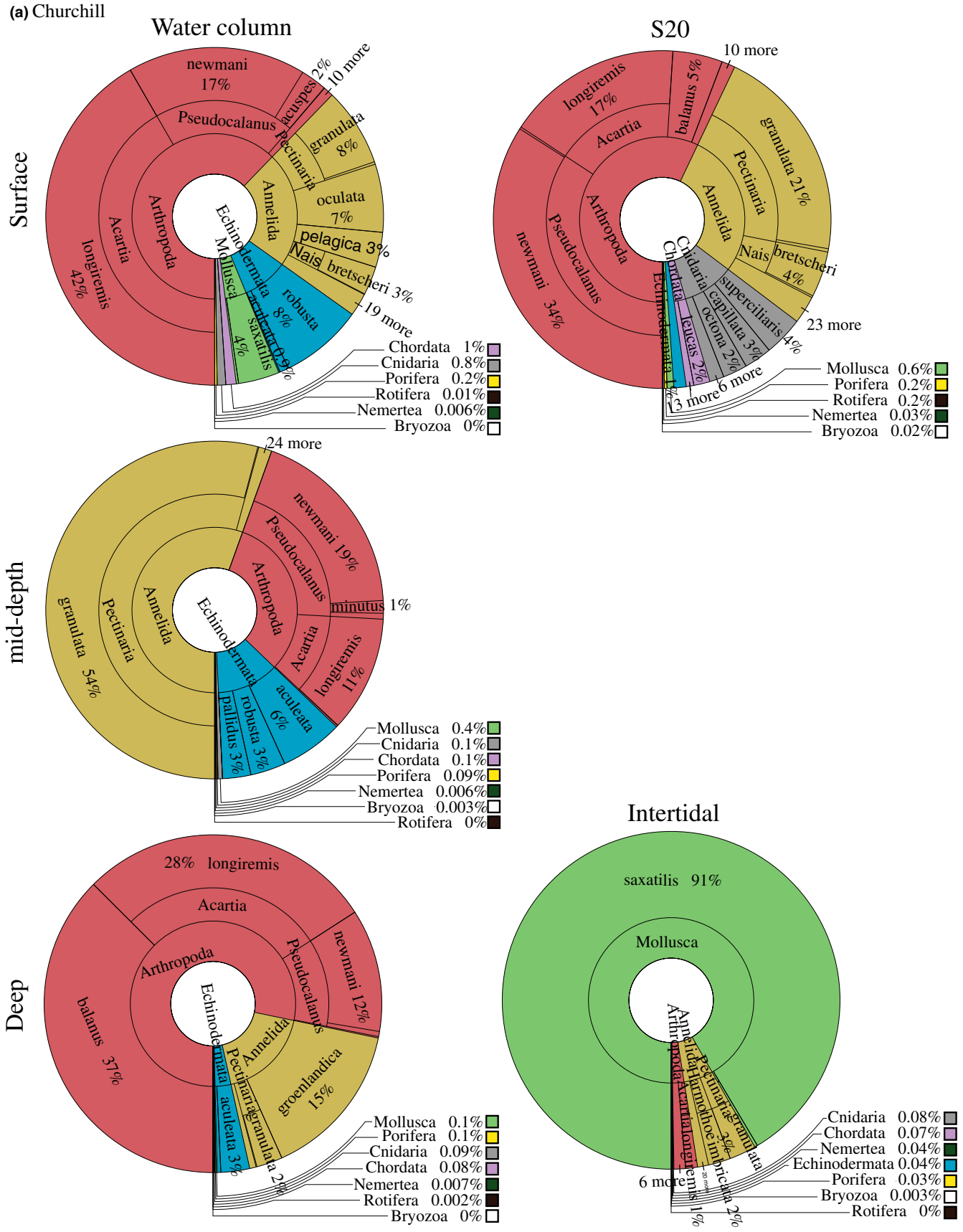


FIGURE 4 eDNA community differences between sampling locations (i.e., water column (surface, middepth and deep), tide pools) and seasons (summer S20 and Fall F20). The different layers represent phyla (central), genus and species (peripheral)

FIGURE 4. continued

column and tide pool samples (averaging 25.40 and 30.27 species in tide pools and water column samples, respectively; $p = 0.42$; Figure 5). In contrast, in Iqaluit, despite the similar number of reads in the tide pool and water column samples (averaging 1,061 and 1,716 reads in tide pools and water column samples, respectively; $p = 0.50$), species richness was significantly greater in tide pools than in the water column (averaging 18.33 and 13.92 species in tide pool and water column samples, respectively; $p = 0.02$; Figure 5). In Iqaluit, the tide pools had estimated Sørenson similarity indices of 0.65, 0.64, 0.62 with the surface, middepth and deep water, respectively, whereas in Churchill, the tide pools had slightly higher estimates of 0.67, 0.84, and 0.68 for the surface, middepth and deep water, respectively.

The community structure differed significantly among the water depths, but the proportion of explained variance was greater for Churchill than Iqaluit (Figure 3, Churchill: $R^2 = 0.13$, $p < 0.001$; Iqaluit: $R^2 = 0.08$, $p = 0.04$), The Crustacean *Balanus balanus* dominated the deep water of both ports (cumulative contributions for Churchill = 80.0% middepth vs. deep water and 67.1% surface vs. deep water; Iqaluit = 62.3% middepth vs. deep water and 65.5% surface vs. deep water) and *Nemertea* was found only in middepth in Iqaluit (Figure 5). In Iqaluit, the Shannon index, species richness and number of reads did not differ significantly among the depth layers (ANOVA shannon: $p = 0.1$; species richness: $p = 0.3$; reads abundance: $p = 0.1$). In contrast, in Churchill, the Shannon index differed significantly among the depth layers ($p \leq 0.001$). Higher species richness was found at the surface ($p = 0.02$), which generally corresponded to where there are more freshwater inputs from the Churchill River (Figure 6). Species detected only at the surface included 52.4% and 19.0% freshwater and brackish species, respectively. The middepth similarity index was the highest among all water depth comparisons (Sørensen and Jaccard nonparametric estimates:

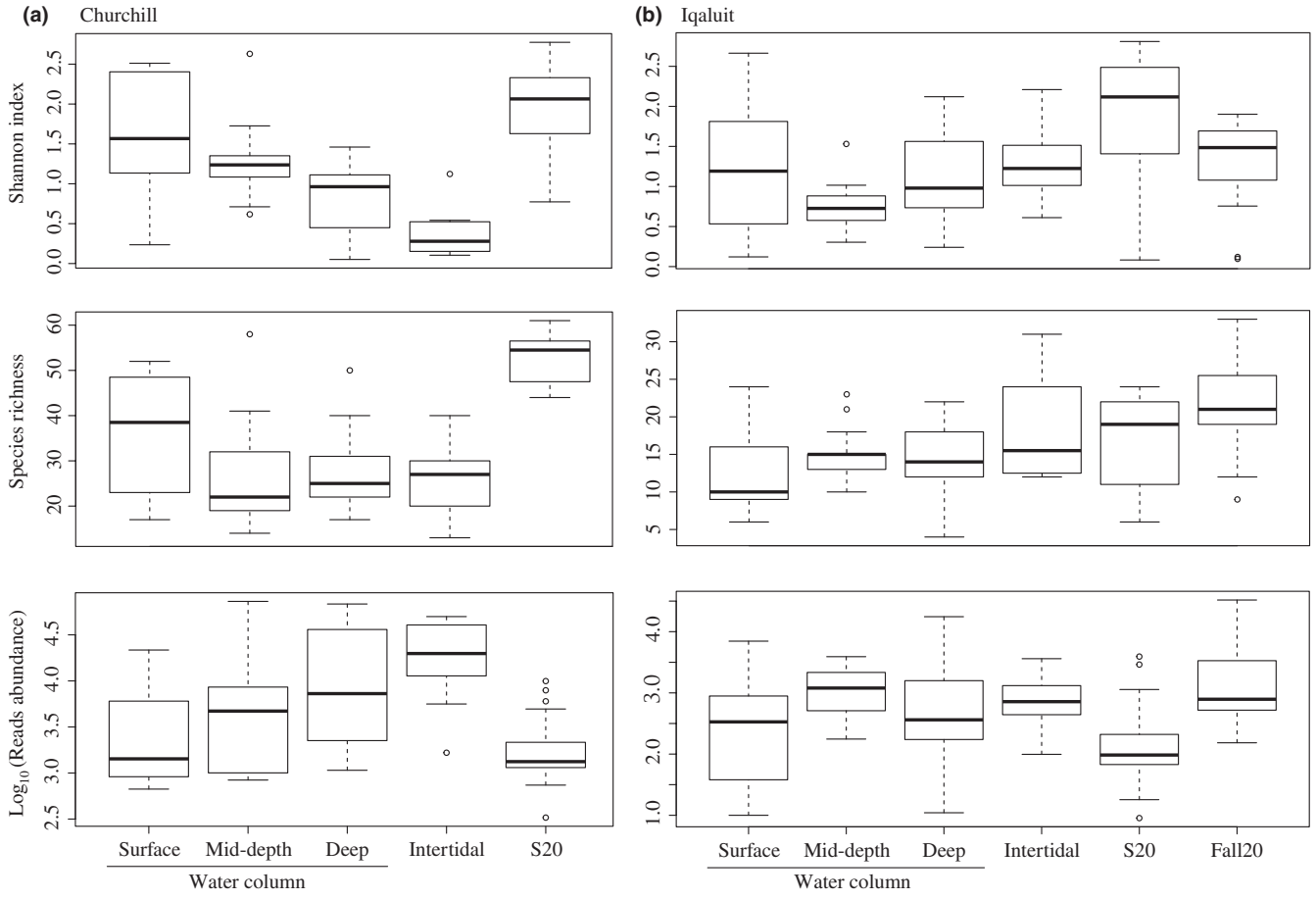


FIGURE 5 Boxplots comparing Shannon indices, species richness, and read abundances detected using eDNA metabarcoding for each sampling location (i.e., water column (surface, middepth and deep), tide pools and S20 and Fall20) in Churchill and Iqaluit. The lines inside the boxes represents the median values, the top and bottom of the boxes represent the 75% and 25% quartiles and outliers are shown using empty circles (i.e., any data beyond 1.5*IQR)

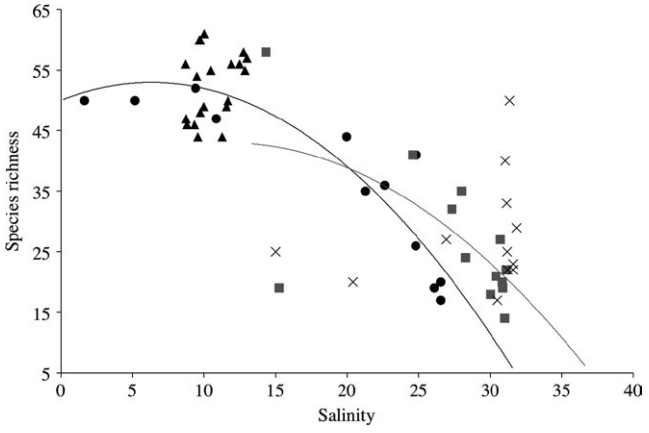


FIGURE 6 Relationship between the species richness detected using eDNA metabarcoding and the salinity of the water collected for the surface layer (R^2 = 0.85, black; circles: sampling water column and S20: triangles) and middepth samples (R^2 = 0.44, gray squares) and deep water (gray cross)

1.0 for Iqaluit and 0.92 for Churchill), but not significantly so relative to the Iqaluit surface-deep and the Churchill intertidal-mid, surface-mid, and surface-deep comparisons.

3.3 | Seasonal variation

The community structure varied significantly between the summer and fall sampling (Figure 3, PERMANOVA; R^2 = 0.30, p < 0.001); Arthropods dominated the summer samples, whereas Annelids dominated in fall (Figure 4) with a total of 54.1% shared species. Species richness was greater under ice cover than in summer (richness: t = 2.3, p = 0.02; Shannon index: t = −2.6, p = 0.01), averaging 21 and 17 species in fall and summer samples, respectively (Figure 5).

4 | DISCUSSION

Improved biodiversity monitoring programs are crucial for maintaining the integrity of coastal marine ecosystems. Evaluating the potential of eDNA to identify Arctic species and understanding the dynamics of eDNA distribution in coastal environments are both timely and important goals for improving biodiversity monitoring. Here, we present evidence that eDNA may be used to assess Arctic biodiversity and show that, despite the cold and well-mixed

environment, standardized eDNA approaches to biodiversity monitoring will need to consider local spatio-temporal variation.

4.1 | Taxonomic assignment challenges

The high congruence between historical Arctic data and eDNA samples (74.0%) supports the efficacy of aquatic eDNA metabarcoding for evaluating Arctic coastal biodiversity at the species level. The species detected with eDNA that were not previously known from the Canadian Arctic (42 species in Churchill and 11 species in Iqaluit) may be new species records, unexpected NIS or Arctic species that are not yet represented in the sequence reference databases that instead matched a closely related non-Arctic species sequence. About 3,894–4,674 (4,284 ± 390) macro- and megabenthic species are estimated to inhabit the Arctic shelf regions (Piepenburg et al., 2011). However, Goldsmit et al. (2014) showed that approximately 15% of the taxa identified in Arctic ports were considered new records within the regions surveyed and approximately 8% within the more extensive adjacent surrounding regions. Piepenburg et al. (2011) suggested that further traditional sampling in the coastal Arctic would increase the number of Mollusca, Arthropoda and Echinodermata species by 26–52%, indicating that between about a fifth and a third of the expected Mollusca-Arthropoda-Echinodermata species pool is still unknown. Given these estimated biases in the historical data, it is therefore not surprising that the congruence between species detected by metabarcoding and historical data is not 100%.

A major shortcoming of metabarcoding is the incomplete state of reference sequence databases. Despite considerable barcoding efforts, reference sequences are still very limited for coastal benthic species, especially for remote regions such as the Arctic. Results showed that ~50% of known Arctic species are actually present in sequence databases and that a similar proportion of the eDNA sequences were assigned to species, indicating that reference database omissions are limiting eDNA metabarcoding surveys at this time and that COI sequencing efforts can rapidly improve Arctic biomonitoring. As shown by the low proportion of OTUs identified at the species level, Porifera and Rotifera were less likely to be detected than other groups such as Annelida (Figure 2). The use of eDNA metabarcoding may thus become a powerful approach to guide reference database improvement (e.g., 97% Rotifera OTUs were not identified at the species level). Moreover, groups such as Bryozoans, Nemerteans and Rotifera are currently not included in the historical Arctic Canada species records that we compiled, but they are important to coastal ecosystems and could be good indicators of biodiversity shifts caused by ice cover changes. The eDNA metabarcoding method might thus be a good practical approach to evaluate the community changes of such species groups, even when poorly identified at the species level. The better our knowledge of local species richness, potential invaders, and their corresponding genetic information, the more accurate our eDNA biodiversity monitoring methods will become. However, even when not assigned to species, the eDNA sequences detected here provide a sequence reference baseline that can be used to evaluate future species loss, new invasions, or other changes in community structure.

Once a taxon has been firmly identified by taxonomic experts and its barcode sequence has been deposited in GenBank or BOLD, eDNA might eventually reduce the need for large teams of expert taxonomists to carry out routine biodiversity monitoring. Yet, the routine application of metabarcoding for Arctic monitoring requires overcoming various limitations. For example, here the eDNA metabarcoding identified *Acartia tonsa*, a potential invader that has been previously recorded in the ecoregions of ports connected to Churchill (Chan et al., 2012). However, the current available COI sequences for *Acartia tonsa* form several distinct clades, some of which cluster with *Acartia hudsonica*, raising the possibility that the eDNA sequences assigned to *A. tonsa* actually belong to the native *A. hudsonica*. Thus, taxonomic expertise remains crucial for reducing biases of species distributions related to increasing use of large-scale eDNA metabarcoding.

Using two COI primer pairs, we increased the level of genetic polymorphism recorded at the species level, thereby increasing the resolution of the method for biodiversity monitoring (Deagle, Jarman, Coissac, Pompanon, & Taberlet, 2014). In addition to increasing the number of species detected, combining multiple primers may also reduce bias of eDNA dominance among species groups (e.g., dominance shift between Arthropoda and Annelida; Figure 2). Despite the fact that the amplification of COI is often desirable to differentiate species using DNA barcoding procedures (Che et al., 2012), the degree of universality for COI primers is relatively low and so combining multiple COI primer pairs as we did enabled monitoring a greater proportion of the diversity. Further studies are, however, needed to evaluate how the combination of the primer sets may depict local species diversity. On the other hand, targeting genes with lower taxonomic specificity (e.g., 18S) could improve the detection of biodiversity shifts at higher levels (e.g., phyla level; see Bik et al., 2012; Deagle et al., 2014; Elbrecht & Leese, 2015).

Characterization of biodiversity with metabarcoding is biased at the amplification step (see Deiner et al., 2017; Freeland, 2017; Kelly et al., 2017 and Pawluczyk et al., 2015). Evaluating the primer bias of eDNA metabarcoding among primer pairs is currently limited due to the unknown nature of eDNA and actual technology used to characterize eDNA. Our selected primer pairs were previously tested on 104 zooplankton species and validated on mock metazoan communities collected in Canadian ports by Zhang (2017). However, even these in situ mock communities are not representative of the complex mixture of eDNA in real biological samples, as they consisted of purified DNA added in equimolar concentrations. Thus, future research evaluating the effects of primer bias is needed. Nevertheless, the results from our current comparisons show that there are important differences in eDNA community composition across space and time in samples

collected using the same sampling and sequencing method. The large number of species detected in this study does allow for establishing a baseline for detecting species from their eDNA and measuring Arctic community structure changes. The current lack of knowledge on primer bias does limit comparisons of species lists and community structure between studies using different primer sets and genetic loci, however.

4.2 | Spatio-temporal eDNA variation

Our results clearly show that metazoan eDNA distribution in Arctic coastal environments has significant temporal and spatial variation. The transport of eDNA may be substantially higher compared to southern regions due to the limited degradation from cold water and the limited UV exposure during much of the year. Although eDNA is expected to be highly dispersed in cold environments, results here show clear horizontal and vertical eDNA heterogeneity in the Arctic. The observed heterogeneity of eDNA within and between samples suggests that, based on the summer and fall sample rarefaction curves, collecting at least 15 samples across as many sites as possible is optimal for comprehensive estimates of biodiversity variation (see Supporting Information Figure S1); an important metric for detecting effects of climate and shipping traffic change. A better understanding of the spatio-temporal variation in eDNA due to local biotic and abiotic conditions will be important in standardizing comparisons of eDNA samples across spatial and temporal gradients in the Arctic marine environment.

Vertical eDNA distribution in the water column may vary as a function of the life cycle of species, transport and settling advection (Turner, Uy, & Everhart, 2015) and complex hydrodynamic processes. In addition to wave action on eDNA mixing (O'Donnell et al., 2017; Port et al., 2016), our data support the idea that in estuarine conditions, such as in Churchill, the freshwater flowing from the river over long distances may contribute to increasing the diversity in the surface water layer (e.g., Deiner & Altermatt, 2014; Jane et al., 2015). Community changes related to eDNA composition thus need to integrate information on temporal variation in river discharge. The variability in the eDNA capture zone should therefore combine complex interactions between community changes and hydrodynamic models.

The dominance of Mollusca reads in tide pools is consistent with the observed species composition in these habitats (e.g., Goldsmit, 2016). However, our results support the hypothesis that tides may modify differences in eDNA composition between the water column and tide pools. At the local scale, the eDNA distribution varied between habitats at both ports (i.e., water column and tide pools), but this pattern was more distinct in Churchill. The large tidal area in Iqaluit increases the water admixture between tide pools and the open ocean (11.72 m maximum tide in Iqaluit and 4.93 m in Churchill (Tide-forecast 2017)), which may explain the relatively lower community differentiation between tide pool and water column sites in Iqaluit compared to Churchill.

Coastal biodiversity monitoring in the Arctic using traditional sampling approaches is generally limited to summer. In contrast to traditional surveys, the quality of eDNA surveys might actually improve under the ice cover due to the limited UV exposure and cold water temperature, hence promoting eDNA preservation and detection (Barnes et al., 2014). On the other hand, cold temperatures are expected to reduce the metabolism of species and associated eDNA release/detection (Lacoursière-Roussel, Rosabal, et al., 2016). Here, eDNA metabarcoding of water collected under ice cover detected greater species richness than summer water collections. This is particularly relevant because the use of eDNA could expand the time window to survey coastal biodiversity in the Arctic. The observed species dominance changes between both seasons may also reflect life history (e.g., late Annelida reproduction; P. Archambault unpublished data). Here our survey is limited to two sampling periods, and thus further studies are needed to differentiate relative effects of species and eDNA ecologies between seasons (Hulbert, 1984).

4.3 | Arctic conservation biology

As contributions of sequences from identified specimens increase to databases such as BOLD, so too will the ability to track biodiversity changes over time at the species level with powerful methods such as eDNA metabarcoding (Gibson et al., 2014; Ji, Ashton, & Pedley, 2013; Taylor & Harris, 2012). In the Arctic, the development of cost-effective monitoring methods is essential for better protecting the integrity of important natural environments and endangered species and to ensure sustainable subsistence harvesting by aboriginal people, as well as recreational and commercial harvest by non-Aboriginals. Applying eDNA metabarcoding to assess biodiversity in remote coastal regions offers several advantages toward increasing the speed and accuracy with which we can amass biodiversity data. As part of this research project, local community members and permanently stationed northern research staff were trained in eDNA sampling techniques with the goal of enabling a network of community-based monitoring. In this context, we optimized eDNA strategies for remote regions. We first used a syringe method for filtering samples (Deiner & Altermatt, 2014), which allows for sampling at multiple sites simultaneously and limits cross-contamination between samples as each sample can be processed with independent equipment. Moreover, the simplicity of this approach allows inexperienced collaborators to collect more eDNA samples per unit of time relative to standard practices of using an electric pump. Second, as storing and shipping frozen samples in remote regions is risky and often not possible, we used methods that allowed for DNA preservation at room temperature (Renshaw, Olds, Jerde, McVeigh, & Lodge, 2014). Lastly, the cost-effective extraction method increases the ability to process large number of samples.

By overcoming methodological issues and improving knowledge about the ecology of eDNA in coastal area, this project

creates the opportunity for future monitoring of metazoan coastal diversity in highly vulnerable ecosystems such as Arctic commercial ports. The combined benefits of being able to identify large numbers of species including local species and potential invaders, assess a large number of phyla, the local habitat variability and together with the effectiveness of the eDNA method under ice cover, are likely to make eDNA metabarcoding an efficient complementary approach to detect large-scale Arctic coastal biodiversity changes. As the eDNA method progresses, the use of eDNA is likely to expand and become a catalyst for increased research on coastal biodiversity, ecosystem services, and sustainability, particularly in remote regions of the world such as the Canadian Arctic. However, spatio-temporal dimensions need to be considered in standardizing and optimizing the assessment of marine biodiversity using eDNA.

ACKNOWLEDGMENTS

This project was funded by ArcticNet, POLAR knowledge, Nunavut Wildlife Management Board and the Fisheries and Oceans Canada Aquatic Invasive Species Monitoring Program. Logistic support was provided by the Churchill Northern Studies Center and Fisheries and Oceans Canada Inuvik area office. We thank Melania Cristescu and Guang Zhang for their help in choosing primers and Frederic Chain and Yiyuan Li for suggestions about the development of the Barque analysis pipeline and data interpretation, Cathryn Abbott for sharing specimens to test in vitro primers, Rob Young for sharing sequences from Arctic species and Jérôme Laroche for his help with the bioinformatics analyses. We are grateful to Brian Boyle, Maëlle Sevellec, Charles Babin, Eric Parent, Jésica Goldsmit, Rob Young, Nathalie Simard, Farrah Chan and Guillaume Côté for their support in various areas of this work. We also thank the following individuals for field assistance, participating in training and assisting with coordination of field logistics: Valérie Cypihot, Frédéric Hartog, Christopher W. Mckindsey, LeeAnn Fishback, Daniel Gibson, Dick Hunter, Austin MacPherson, Kimberly Thomson, Heather Clark, Colin Gallagher, Zoya Martin, Chris Lewis, Jeremiah Young, Richard Moore, Thomas Whittle, Rory McDonald, Frederic Lemire, Levi tikivik, Fred Lemire, Conner McCormack and students from the Nunavut Arctic college. We are also grateful to the Managing Editor Karen Chambers, the Subject Editor Aurélie Bonin and three anonymous referees for their constructive comments on a previous version of the manuscript.

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

Anaïs Lacoursière-Roussel is a conservation ecologist evaluating anthropogenic impacts on the dynamic of aquatic communities. All authors of this manuscript are interested developing and calibrating the eDNA method in the application of aquatic species distribution to improve the efficiency of conservation planning. ALR, LB, KH, PA, EG, and DL conceived the idea, ALR, LB, KH, KD, PA, and EN structured and edited the manuscript, KH and ALR developed the study design and participated in field collections. KH and PA are specialized in the Arctic coastal surveillance, CH, ALR, KD, and NL developed laboratory methodology and EN developed the analysis pipeline.

DATA ACCESSIBILITY

Raw sequences reads were deposited in NCBI's Sequence Read Archive (SRA, http://www.ncbi.nlm.nih.gov/sra) under Bioproject PRJNA388333.

ORCID

Anaïs Lacoursière-Roussel http://orcid.org/0000-0002-9345-5682

REFERENCES

ACIA (2004). *Impacts of a warming Arctic: Arctic climate impact assessment* (146 pp). Cambridge, UK: Cambridge University Press.

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, *26*, 32–46.

Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*. Babraham Bioinformatics. Retrieved from www.bioinformatics.babraham.ac.uk/projects/fastqc

Archambault, P., Snelgrove, P. V. R., Fisher, J. A. D., Gagnon, J.-M., Garbary, D. J., Harvey, M., … Poulin, M. (2010). From sea to sea: Canada's three oceans of biodiversity. *PLoS ONE*, *5*, e12182. https://doi.org/10.1371/journal.pone.0012182

Arctic Council (2009). *Arctic marine shipping assessment 2009 report Tromso*. Akureykri: PAME, 194 pages

Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, *17*, 1–17. https://doi.org/10.1007/s10592-015-0775-4

Barnes, M. A., Turner, C. R., Jerde, C. L., Renshaw, M. A., Lindsay Chadderton, W., & Lodge, D. M. (2014). Environmental conditions influence eDNA persistence in aquatic systems. *Environmental Science and Technology*, *48*, 1819–1827. https://doi.org/10.1021/es404734p

Bik, H. M., Porazinska, D. L., Creer, S., Caporaso, J. G., Knight, R., & Thomas, W. K. (2012). Sequencing our way towards understanding global eukaryotic biodiversity. *Trends in Ecology and Evolution*, *27*, 233–243. https://doi.org/10.1016/j.tree.2011.11.010

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics*, *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Chain, F. J. J., Brown, E. A., MacIsaac, H. J., & Cristescu, M. E. (2016). Metabarcoding reveals strong spatial structure and temporal turnover of zooplankton communities among marine and freshwater ports. *Diversity and Distributions*, *22*, 493–504. https://doi.org/10.1111/ddi.12427

Chan, F. T., Bailey, S. A., Wiley, C. J., & MacIsaac, H. J. (2013). Relative risk assessment for ballast-mediated invasions at Canadian Arctic ports. *Biological Invasions*, *15*, 295–308. https://doi.org/10.1007/s10530-012-0284-z

Chan, F. T., Bronnenhuber, J. E., Bradie, J. N., Howland, K. L., Simard, N., Bailey, S. A., & Department of Fisheries and Oceans, Ottawa, ON(Canada); Canadian Science Advisory Secretariat, Ottawa, ON (Canada). (2012). *Risk assessment for ship-mediated introductions of aquatic nonindigenous species to the canadian arctic. Canadian science advisory secretariat*. Research Document 2011/105. Ottawa, ON: DFO.

Chan, F. T., MacIsaac, H. J., & Bailey, S. A. (2015). Relative importance of vessel hull fouling and ballast water as transport vectors of nonindigenous species to the Canadian Arctic. *Journal canadien des sciences halieutiques et aquatiques*, *72*, 1230–1242. https://doi.org/10.1139/cjfas-2014-0473

Chao, A., Ma, K. H., Hsieh, T. C., & Chiu, C.-H. (2016). *SpadeR: Species-richness prediction and diversity estimation with R*. R package version 0.1.1. Retrieved from https://cran.r-project.org/web/packages/SpadeR

Che, J., Chen, H.-M., Yang, J.-X., Jin, J.-Q., Jiang, K., Yuan, Z.-Y., … Zhang, Y.-P. (2012). Universal COI primers forDNA barcoding amphibians. *Molecular Ecology Resources*, *12*, 247–258. https://doi.org/10.1111/j.1755-0998.2011.03090.x

Cheung, W. W. L., Lam, V. W. Y., Sarmiento, J. L., Kearney, K., Watson, R., & Pauly, D. (2009). Projecting global marine biodiversity impacts under climate change scenarios. *Fish and Fisheries*, *10*, 235–251. https://doi.org/10.1111/j.1467-2979.2008.00315.x

Cusson, M., Archambault, P., & Aitken, A. (2007). Biodiversity of benthic assemblages on the Arctic continental shelf: Historical data from Canada. *Marine Ecology Progress Series*, *331*, 291–304. https://doi.org/10.3354/meps331291

Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F., & Taberlet, P. (2014). DNA metabarcoding and the cytochrome c oxidase subunit I marker: Not a perfect match. *Biology Letters*, *10*, 20140562. https://doi.org/10.1098/rsbl.2014.0562

Deiner, K., & Altermatt, F. (2014). Transport distance of invertebrate environmental DNA in a natural river. *PLoS ONE*, *9*, e88786. https://doi.org/10.1371/journal.pone.0088786

Deiner, K., Bik, H., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., … Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, *26*, 5872–5895. https://doi.org/10.1111/mec.14350

Eguíluz, V. M., Fernández-Gracia, J., Irigoien, X., & Duarte, C. M. (2016). A quantitative assessment of Arctic shipping in 2010–2014. *Scientific Reports*, *6*, 30682. https://doi.org/10.1038/srep30682

Elbrecht, V., & Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass—sequence relationships with an innovative metabarcoding protocol. *PLoS ONE*, *10*, e0130324. https://doi.org/10.1371/journal.pone.0130324

Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome C oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, *3*, 294–299.

Foote, A. D., Thomsen, P. F., Sveegaard, S., Wahlberg, M., Kielgast, J., Kyhn, L. A., … Gilbert, T. P. (2012). Investigating the potential use of environmental DNA (eDNA) for genetic monitoring of marine mammals. *PLoS ONE*, *7*, e41781. https://doi.org/10.1371/journal.pone.0041781

Freeland, J. R. (2017). The importance of molecular markers and primer design when characterizing biodiversity from environmental DNA. *Genome*, *60*, 358–374. https://doi.org/10.1139/gen-2016-0100

Gavrilchuk, K., & Lesage, V. (2014). Large-scale marine development projects (mineral, oil and gas, infrastructure) proposed for Canada's North. *Canadian Technical Report of Fisheries and Aquatic Sciences*, *3069*, viii-84.

Geller, J. B., Meyer, C. P., Parker, M., & Hawk, H. (2013). Redesign of PCR primers for mitochondrial Cytochrome c oxidase subunit I for marine invertebrates. *Molecular Ecology Resources*, *13*, 851–861. https://doi.org/10.1111/1755-0998.12138

Gibson, J., Shokralla, S., Porter, T. M., King, I., van Konynenburg, S., Janzen, D. H., … Hajibabaei, M. (2014). Simultaneous assessment of the macrobiome and microbiome in a bulk sample of tropical arthropods through DNA metasystematics. *Proceedings of the National Academy of Sciences*, *111*, 8007–8012. https://doi.org/10.1073/pnas.1406468111

Goldsmit, J. (2016). *Benthic non-indigenous species in ports of the Canadian Arctic: identification, biodiversity and relationship with global warming and shipping activity*. ISMER-Université du Québec à Rimouski.

Goldsmit, J., Archambault, P., Chust, G., Villarino, E., Liu, G., Lukovich, J. V., … Howland, K. L. (2018). Projecting present and future habitat suitability of ship-mediated aquatic invasive species in the Canadian Arctic. *Biological Invasions*, *20(2)*, 501–517.

Goldsmit, J., Howland, K. L., & Archambault, P. (2014). Establishing a baseline for early detection of non-indigenous species in ports of the Canadian Arctic. *Aquatic Invasions*, *9*, 327–342. https://doi.org/10.3391/ai

Gotelli, N. J., & Colwell, R. K. (2010). Estimating species richness. In *Biological diversity: Frontiers in measurement and assessment* (p. 345). Oxford, UK: Oxford University Press.

Grosholz, E. (2002). Ecological and evolutionary consequences of coastal invasions. *Trends in Ecology and Evolution*, *17*, 22–27. https://doi.org/10.1016/S0169-5347(01)02358-8

Guy, E., & Lasserre, F. (2016). Commercial shipping in the Arctic: New perpectives, challenges and regulations. *Polar Record*, *52*, 294–304. https://doi.org/10.1017/S0032247415001011

Hulbert, R. C. (1984). Paleoecology and population dynamics of the early Miocene (Hemingfordian) horse *Parahippus leonensis* from the Thomas Farm site, Florida. *Journal of Vertebrate Paleontology*, *4*, 547–558. https://doi.org/10.1080/02724634.1984.10012030

Jane, S. F., Wilcox, T. M., McKelvey, K. S., Young, M. K., Schwartz, M. K., Lowe, W. H., … Whiteley, A. R. (2015). Distance, flow, and PCR inhibition: eDNA dynamics in two headwater streams. *Molecular Ecology Resources*, *15*, 216–227. https://doi.org/10.1111/1755-0998.12285

Ji, Y., Ashton, L., & Pedley, S. M. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, *16*, 1245–1257. https://doi.org/10.1111/ele.12162

Kelly, R. P., Closek, C. J., O'Donnell, J. L., Kralj, J. E., Shelton, A. O., & Samhourri, J. F. (2017). Genetic and manual survey methods yield different and complementary views of an ecosystem. *Frontiers in Marine Science*, *3*, 1–11.

Lacoursière-Roussel, A., Côté, G., Leclerc, V., & Bernatchez, L. (2016). Quantifying relative fish abundance with eDNA: A promising tool for fisheries management. *Journal of Applied Ecology*, *53*, 1148–1157. https://doi.org/10.1111/1365-2664.12598

Lacoursière-Roussel, A., Dubois, Y., & Bernatchez, L. (2016). Improving herpetological surveys in eastern North America using the environmental DNA method. *Genome Special Edition*, *59*, 991–1007.

Lacoursière-Roussel, A., Rosabal, M., & Bernatchez, L. (2016). Estimating fish abundance and biomass from eDNA concentrations: Variability among capture methods and environmental conditions. *Molecular Ecology Resources*, *16*, 1401–1414. https://doi.org/10.1111/1755-0998.12522

Larigauderie, A., Prieur- Richard, A.-H., Mace, G. M., Lonsdale, M., Mooney, H. A., Brussaard, L., … S., T. (2012). Biodiversity and ecosystem services science for a sustainable planet: The DIVERSITAS vision for 2012–20. *Current Opinion in Environnemtal Sustainability*, *4*, 101–105. https://doi.org/10.1016/j.cosust.2012.01.007

Legendre, P., & Legendre, L. (1998). Numerical ecology. In *Developments in environmental modelling* (2nd ed., 20, pp. xv + 853), New York, NY: Elsevier.

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., … S., R. J. (2013). A new versatile primer set targeting a

short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, *10*, e34. https://doi.org/10.1186/1742-9994-10-34

Lim, N. K. M., Tay, Y. C., Srivathsan, A., Tan, J. W. T., Kwik, J. T. B., Baloğlu, B., … S., D. C. J. (2016). Next-generation freshwater bioassessment: eDNA metabarcoding with a conserved metazoan primer reveals species-rich and reservoir-specific communities. *Royal Society Open Science*, *3*, 160635. https://doi.org/10.1098/rsos.160635

Link, H., Piepenburg, D., & Archambault, P. (2013). Are Hotspots Always Hotspots? The relationship between diversity, resource and ecosystem functions in the Arctic. *PLoS ONE*, *8*, e74077. https://doi.org/10.1371/journal.pone.0074077

Lodge, D. M., Turner, C. R., Jerde, C. L., Barnes, M. A., Chadderton, L., Egan, S. P., … Pfrender, M. E. (2012). Conservation in a cup of water: Estimating biodiversity and population abundance from environmental DNA. *Molecular Ecology*, *21*, 2555–2558. https://doi.org/10.1111/j.1365-294X.2012.05600.x

López, E., Olivier, F., Grant, C., & Archambault, P. (2016). A new species and four new records of sedentary polychaetes from the Canadian High Arctic. *Journal of the Marine Biological Association of the United Kingdom*, *97*, 1685–1694.

Magoč, T., & Salzberg, S. L. (2011). FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, *27*, 2957–2963.

Mahé, F., Rognes, T., Quince, C., De Vargas, C., & Dunthorn, M. (2015). Swarm v2: Highly scalable and high-resolution amplicon clustering. *PeerJ*, *3*, e1420. https://doi.org/10.7717/peerj.1420

Mahon, A. R., Jerde, C. L., Galaska, M., Bergner, J. L., Chadderton, W. L., Lodge, D. M., … S., L. G. (2013). Validation of eDNA surveillance sensitivity for detection of Asian carps in controlled and field experiments. *PLoS ONE*, *8*, e58316. https://doi.org/10.1371/journal.pone.0058316

Miller, A. W., & Ruiz, G. M. (2014). Arctic shipping and marine invaders. *Nature Climate Change*, *4*, 413–416. https://doi.org/10.1038/nclimate2244

O'Donnell, J. L., Kelly, R. P., Shelton, A. O., Samhouri, J. F., Lowell, N. C., & Williams, G. D. (2017). Spatial distribution of environmental DNA in a nearshore marine habitat. *PeerJ*, *5*, e3044. https://doi.org/10.7717/peerj.3044

Olivier, F., San Martín, G., & Archambault, P. (2013). Revision of Streptospinigera Kudenov, 1983 (*Polychaeta*, *Syllidae*, *Anoplosyllinae*) and description of a new species from the Arctic and the western northern Atlantic. *Polar Biology*, *36*, 1499–1507. https://doi.org/10.1007/s00300-013-1369-6

Parker, I. M., Simberloff, D., Lonsdale, W. M., Goodell, K., Wonham, M., Kareiva, P. M., … Goldwasser, L. (1999). Impact: Toward a framework for understanding the ecological effects of invaders. *Biological Invasions*, *1*, 3–19. https://doi.org/10.1023/A:1010034312781

Pawluczyk, M., Weiss, J., Links, M. G., Aranguren, M. E., Wilkinson, M. D., & Egea-Cortines, M. (2015). Quantitative evaluation of bias in PCR amplification and next-generation sequencing derived from metabarcoding samples. *Analytical and Bioanalytical Chemistry*, *407*, 1841–1848. https://doi.org/10.1007/s00216-014-8435-y

Piepenburg, D., Archambault, P., Ambrose, W. G., Blanchard, A. L., Bluhm, B. A., Carroll, M. L., … Włodarska-Kowalczuk, M. (2011). Towards a pan-Arctic inventory of the species diversity of the macro- and megabenthic fauna of the Arctic shelf seas. *Marine Biodiversity*, *41*, 51–70. https://doi.org/10.1007/s12526-010-0059-7

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R Core Team (2017). *nlme: Linear and nonlinear mixed effects models*. R package version 3.1-131, Retrieved fromhttps://CRAN.R-project.org/package=nlme

Port, J. A., O'Donnell, J. L., Romero-Maraccini, O. C., Leary, P. R., Litvin, S. Y., Nickols, K. J., … Kelly, R. P. (2016). Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. *Molecular Ecology*, *25*, 527–541. https://doi.org/10.1111/mec.13481

Ratnasingham, S., & Hebert, P. D. N. (2007). bold: The barcode of life data system (http://www.barcodinglife.org). *Molecular Ecology Notes*, *7*, 355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x

Renshaw, M. A., Olds, B. P., Jerde, C. L., McVeigh, M. M., & Lodge, D. M. (2014). The room temperature preservation of filtered environmental DNA samples and assimilation into a phenol–chloroform–isoamyl alcohol DNA extraction. *Molecular Ecology Resources*, *15*, 168–176.

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, *4*, e2584. https://doi.org/10.7717/peerj.2584

Roy, R., Iken, K., & Archambault, P. (2015). Regional variability of megabenthic community structure across the Canadian Arctic. *Arctic*, *68*, 180–192. https://doi.org/10.14430/arctic4486

Sassoubre, L. M., Yamahara, K. M., Gardner, L. D., Block, B. A., & Boehm, A. B. (2016). Quantification of Environmental DNA (eDNA) shedding and decay rates for three marine fish environ. *Environmental Science and Technology*, *50*, 10456–10464. https://doi.org/10.1021/acs.est.6b03114

Shokralla, S., Porter, T. M., Gibson, J. F., Dobosz, R., Janzen, D. H., Hallwachs, W., … Hajibabaei, M. (2015). Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform. *Scientific Reports*, *5*, 9687. https://doi.org/10.1038/srep09687

Smith, L. R., & Stephenson, S. R. (2013). New Trans-Arctic shipping routes navigable by midcentury. *Proceedings of the National Academy of Sciences*, *11*, 4871–4872.

Taberlet, P., Coissac, E., Hajibabaei, M., & Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology*, *21*, 1789–1793. https://doi.org/10.1111/j.1365-294X.2012.05542.x

Taylor, H. R., & Harris, W. E. (2012). An emergent science on the brink of irrelevance: A review of the past 8 years of DNA barcoding. *Molecular Ecology Resources*, *12*, 377–388. https://doi.org/10.1111/j.1755-0998.2012.03119.x

Thomsen, P. F., Kielgast, J., Iversen, L. L., Møller, P. R., Rasmussen, M., & Willerslev, E. (2012). Detection of a diverse marine fish fauna using environmental DNA from seawater samples. *PLoS ONE*, *7*, e41732. https://doi.org/10.1371/journal.pone.0041732

Thomsen, P., Kielgast, J. O. S., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., … Willerslev, E.. (2012). Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular Ecology*, *21*, 2565–2573. https://doi.org/10.1111/j.1365-294X.2011.05418.x

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, *183*, 4–18. https://doi.org/10.1016/j.biocon.2014.11.019

Tide-forecast (2017). *Tide time & Tide charts for the world*. Retrieved from https://www.tide-forecast.com/locations.

Turner, C. R., Uy, L. K., & Everhart, R. C. (2015). Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biological Conservation*, *183*, 93–102. https://doi.org/10.1016/j.biocon.2014.11.017

Wegleitner, B. J., Jerde, C. L., Tucker, A., Chadderton, W. L., & Mahon, A. R. (2015). Long duration, room temperature preservation of filtered eDNA samples. *Conservation Genetics Resources*, *7*, 789–791. https://doi.org/10.1007/s12686-015-0483-x

Young, R. G., Abbott, C. L., Therriault, T. W., & Adamowicz, S. J. (2016). Barcode-based species delimitation in the marine realm: A test using *Hexanauplia* (*Multicrustacea*: *Thecostraca* and *Copepoda*). *Genome*, *60*, 169–182.

Zhang, G. K. (2017). *Calibration of the multi-gene metabarcoding approach as an efficient and accurate biomonitoring tool*. Master Thesis. McGill University, Québec (Qc) Canada. 90

p. Retrieved from http://digitool.library.mcgill.ca/webclient/StreamGate?folder_id=0&dvs=1523467293609~14

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Lacoursière-Roussel A, Howland K, Normandeau E, et al. eDNA metabarcoding as a new surveillance approach for coastal Arctic biodiversity. *Ecol Evol*. 2018;00:1–15. https://doi.org/10.1002/ece3.4213

Supplementary files

eDNA metabarcoding as a new surveillance approach for coastal Arctic biodiversity

Anaïs Lacoursière-Roussel[1], Kimberly Howland[2], Eric Normandeau[3], Erin Grey[4], Philippe Archambault[5], Kristy Deiner[6], David M. Lodge[7], Cecilia Hernandez[3], Noémie Leduc[3] and Louis Bernatchez[3]

[1] Fisheries and Oceans Canada, St. Andrews Biological Station (SABS), St. Andrews (New Brunswick), Canada, E5B 2L9

[2] Fisheries and Oceans Canada, Central and Arctic Region, Freshwater Institute, Winnipeg (Manitoba), Canada, R3T 2N6,

[3] Université Laval, Institut de Biologie Intégrative et des Systèmes (IBIS), Department of Biology, Québec (Québec), Canada, G1V 0A6

[4] Governors State University, Division of Science, Mathematics and Technology, University Park (Illinois), USA, 60484

[5] Université Laval, Québec-Océan, Department of Biology, Québec (Québec), Canada, G1V 0A6

[6] Department of Evolutionary Biology and Environmental Studies, University of Zurich, 13 Winterthurerstr. 190, CH-8057 Zürich, Switzerland

[7] Cornell University, Department of Ecology and Evolutionary Biology, Ithaca (New York), USA, 14853

Appendix S1. eDNA extraction, amplification and sequencing

To isolate and purify eDNA, 30 µl of Proteinase K (Qiagen) was added to the tubes containing the filter and the Longmire lysis buffer. Tubes were vortexed and incubated at 55°C overnight. After incubation, the filter and lysis buffer mixture was centrifuged one minute at 13,000 RPM in a QIAshredder tube. 950 µl of the solution was then transferred to a new tube and 950 µl of the organic phase of phenol chloroform isoamyl alcohol (i.e. PCI, 25:24:1, Sigma P2069) was added. Tubes were hand shaken for five minutes and centrifuged for five minutes at 10,000 RPM. Supernatant was removed into a new tube and 950 µl of Chloroform-Isoamyl alcohol (CI, 24:1) was added to each tube. Tubes were then shaken for five minutes and centrifuged for five minutes at 10,000 RPM. 750 µl of the supernatant was transferred into a new tube and 750 µl of ice cold isopropanol and 375 µl of room temperature 5M NaCl were added to each tube and left overnight at -4ºC. Tubes were centrifuged for 20 minutes at 13,000 RPM and isopropanol was then carefully poured off. 1,500 µl of cold Ethanol 70% was added and centrifuged for 20 minutes at 13,000 RPM. Ethanol was then carefully poured off and tubes were air dried with lid open in a laminar flow hood for 15 minutes. DNA was resuspended in 80 µl sterilized water (diH_20), placed in an incubator at 55°C for ten minutes and at 4°C overnight to resuspend the DNA. The extracted DNA was then frozen at -20 °C until amplification.

DNA amplifications were performed in a one-step dual-indexed PCR approach specifically designed for Illumina instruments by the "Plate-forme d'Analyses Génomiques" (IBIS, Université Laval). Two pairs of universal mitochondrial cytochrome c oxidase subunit I (COI) primers were used to amplify eDNA from as many metazoan taxa as possible: the forward mlCOIintF (Leray *et al.* 2013) and reverse jgHCO2198 (Geller et al. 2013) amplifying 313bp (hereafter called COI1) and the forward LCO1490 (Folmer *et al.* 1994) and reverse ill_C_R (Shokrella *et al.* 2015) amplifying 367bp (COI2). The primers were tailed on the 5' end with part of the Illumina Nextera adaptors. The following adaptor sequence (regions that anneal to the flowcell and library specific barcodes) and oligonucleotide sequences were used for amplification: AATGATACGGCGACCACCGAGATCTACAC-[INDEX]-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-[Forward primers] and reverse primers CAAGCAGAAGACGGCATACGAGAT-[INDEX]-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-[reverse primers]. Please note that primers used in this work contain Illumina specific sequences protected by intellectual property (Oligonucleotide sequences © 2007-2013 Illumina, Inc. All rights reserved. Derivative works created by Illumina customers are authorized for use with Illumina instruments and products only. All other uses are strictly prohibited).

Three PCR replicates were done for each sample and each primer set. The final reaction volume for each PCR replicate was 24 µL; including 12.5. µl Qiagen Multiplex Mastermix, 6.5 µl diH_20, 1 µl of each primer (10µM), and 3.0 µL of DNA. For all samples, the PCR mixture was denatured at 95°C for 15 min, followed by 35 cycles (94°C for 30 s, 54°C for 90 s (except for the COI2 primers, which was at 52°C for 90 s) and 72°C for 60s) and a final elongation at 72°C for 10 min. Products of the three

aliquots were pooled for each sample. Because barcodes were different for each sample, a negative PCR control was done for each sample and primer set. All amplifications were visualized on a 1.5% agarose gel electrophoresis. If positive amplification of the PCR negative control was observed, amplification was redone with a newly diluted primer set. Because PCRs negative controls had the same barcode as the samples we could not pool them for sequencing. In contrast, field and extraction negative controls were treated exactly as regular samples and were independently indexed and pooled for sequencing. Pooled products were purified using Axygen PCR clean up kit following the manufacture's recommended protocol. Libraries were quantified by AccuClear Ultra High Sensitivity dsDNA Quantitation Kit using the TECAN Spark 10M Reader for each sample and samples were pooled in equal molar concentrations to maximize equal sequence depth per sampling location (150 and 37 ng per site for COI1 and COI2 primer sets respectively in Churchill and 200 and 300 ng for COI1 and COI2 primer sets respectively in Iqaluit). When Quant-iT PicoGreen (Life Technologies) did not detect any DNA, 22.0 µL PCR mixtures were mixed nonetheless (see Table S3 for the concentration and volume for each sample separately).

Sequencing was carried out using an Illumina MiSeq (Illumina, San Diego, USA) at IBIS using a paired-end MiSeq Reagent Kit V3 (Illumina, San Diego, USA; sequence length = 300bp) and following the manufacturer's instructions. For sequencing, the amplicon pool was diluted to 4 nM with molecular grade water, denatured and then sequenced at 10 pM following manufacturer's instructions inclusive of spiking the samples with 15% of PhiX. Adaptor sequence and primer sequences were removed and raw sequencing reads de-multiplexed among samples using the MiSeq Control software v 2.3 into independent libraries. De-multiplexed reads were provided in gzip compressed Fastq format.

Table S1. *In vitro* validation of primers. Native species and potential invaders (low and high risk based on Chan et al. (2012), Chan et al. (2015) and Goldsmit (2016)) that primers were tested *in vitro*. '1' depict a positive PCR amplification and '0' that no amplification occurred.

Species	Phylum	COI	
		COI1 (mlCOIintF-jgHCO2198)	COI2(LCO149-illCR)
Native			
Gammarus oceanicus	Arthropoda	1	1
Astarte elliptica complexe	Mollusca	0	0
Musculus discors	Mollusca	0	0
Macoma calcarea	Mollusca	1	0
Hiatella arctica	Mollusca	0	1
Testudinalia testudinalis	Mollusca	1	0
Margarites groenlandicus	Mollusca	0	1
Margarites helicinus	Mollusca	1	0
Littorina saxatilis	Mollusca	1	1
Littorina obtusata	Mollusca	1	0
Macoma balthica	Mollusca	1	0
Mya truncata	Mollusca	1	0
Potential invader			
Crassostrea gigas	Mollusca	0	0
Botryllus schlosseri	Chordata (Tunicata)	0	1
Ciona intestinalis	Chordata (Tunicata)	1	1
Styela clava	Chordata (Tunicata)	0	0
Jassa marmorata	Arthropoda (amphipoda)	1	1
Crepidula fornicata	Mollusca	1	0
Highly-potential invader			
Caprella mutica	Arthropoda (amphipoda)	1	1
Littorina littorea	Mollusca	1	1
Botrylloides violaceus	Tunicate	0	1
Carcinus maenas	Arthropoda (decapoda)	1	0
Mya arenaria	Mollusca	1	0

Appendix S2. References for native Arctic metazoans list include within the Fisheries and Oceans Canada Arctic Marine Invertebrate Database.

Benthic Invertebrate References

1. Aitken AE, Gilbert R (1986) The biota of intertidal flats at Pangnirtung Fiord, Baffin Island, Northwest Territories. *Naturaliste can* **113**, 191–200.
2. Aitken AE, Risk MJ, Howard JD (1988) Animal-sediment relationships on a subarctic intertidal flat, Pangnirtung Fiord, Baffin Island, Canada. *Journal of Sedimentary Petrology* **58**, 969–978.
3. Atkinson EG, Wacasey JW (1989) Benthic invertebrates collected from Hudson Bay, Canada, 1953 to 1965. *Canadian Data Report of Fisheries and Aquatic Sciences* **744**, iv + 121 p.
4. Cusson M, Archambault P, Aitken A (2007) Biodiversity of benthic assemblages on the Arctic continental shelf: historical data from Canada. *Marine Ecology Progress Series* **331**, 291–304.
5. Ellis DV (1957) Marine infaunal benthos in Arctic North America. PhD thesis for the Faculty of graduate studies and research, Zoology, McGill University, Montreal, Canada.
6. Ellis DV (1960) Marine infaunal benthos in Arctic North America. *Arctic Institute of North America, Technical Paper N°5.*
7. Hopky GE, Lawrence MJ, McRae SM, Chiperzak DB (1994) List of Scientific names of Algae, Invertebrates and Vertebrates Captured under NOGAP Subprojects B.2.1 and B.2.3. 1984–1988. *Canadian Data Report of Fisheries and Aquatic Sciences* **924**, 76 pp.
8. ITIS (Integrated Taxonomic Taxonomic System) Available at http://www.itis.gov/.
9. Jirkov IA, Leontovich MK (2012) Biogeography of Polychaeta of the Aurasian North Polar Basin. *Invertebrate Zoology* **9**, 41–51.
10. Natural Museum. From P. Pocklington report.
11. North South Consultants (2006) Potential dispersal of aquatic invasive species into Hudson Bay from ballast water from ships travelling from ports in Europe and North America.
12. OBIS (Ocean Biogeographic Information system) Available at www.iobis.org.
13. Stewart PL, Pocklington P, Cunjak RA (1985) Distribution, Abundance and Diversity of Benthic Macroinvertebrates on the Canadian Continental Shelf and Slope of Southern Davis Strait and Ungava Bay. *Arctic* **36**, 281–291.
14. Thomson DH (1982) Marine benthos in the Eastern Canadian High Arctic: Multivariate analyses of standing crop and community structure. *Arctic* **35**, 61–74.
15. Thomson DH, Martin CM, Cross WE (1986) Identification and characterization of Arctic nearshore benthic habitats. *Canadian technical report of Fisheries and Aquatic Science N°1434.*

16. Samuelson GM (2001) Polychaetes as indicators of environmental disturbance on subarctic tidal flats, Iqaluit, Baffin Island, Nunavut Territory. *Marine Pollution Bulletin* **42**, 733–741.
17. SLB (Sea Life Base) Available at http://www.sealifebase.org/.
18. Stewart DB (2011) Species inhabiting the Tarium Niryutait Marine Protected Areas in the Canadian Beaufort Sea – Mackenzie Delta.
19. Wacasey JW, Atkinson EG, Glasspool L (1979) Zoobenthos data from upper Frobisher Bay, 1967–1973. *Canadian Data Report of Fisheries and Aquatic Sciences* **164**, 99.
20. Wacasey JW, Atkison EG, Glasspoole L, Bédard C (1980a) A marine biological study of Breevort Harbour and nearby waters of Eastern Baffin Island – Zoobenthos. *Canadian manuscript report of Fisheries and Aquatic Science N° 1557.*
21. Wacasey JW, Atkinson EG, Glasspoole L (1980b) Zoobenthos data from inshore stations of upper Frobisher Bay, 1969-1976. *Canadian Data Report of Fisheries and Aquatic Sciences* **205,** 42.
22. WoRMS (World Register of Marine Species) Available at www.marinespecies.org.

Zooplankton References

1. Andersen OGN (1981) Redescription of Marrus orthocanna (Kramp, 1942) (Cnidaria, Siphonophora). *Steenstrupia* **7**, 293–307.
2. Andronov VN, Kosobokova KN (2011) New species of small, bathypelagic calanoid copepods from the Arctic Ocean: Brodskius arcticus sp. nov. (Tharybidae) and three new species of Pertsovius gen. nov. (Discoidae). *Zootaxa* **2809**, 33–46.
3. Angel MV (1976) Bathyconchoecia arctica n. sp., a new species of ostracod (Halocyprididae, Myodocopidae) from the Arctic. *Crustaceana* **31**, 59–65.
4. ARCOD database. Available at http://www.arcodiv.org.
5. Audzijonyte A Väinölä R (2005) Diversity and distributions of circumpolar fresh- and brackish-water Mysis (Crustacea: Mysida): descriptions of M. relicta Lovén, 1862, M. salemaai n.sp., M. segerstralei n.sp. and M. diluviana n.sp., based on molecular and morphological characters. *Hydrobiologia* **544**, 88–141.
6. Auel H, Hagen W (2002) Mesozooplankton community structure, abundance and biomass in the central Arctic Ocean. *Marine Biology Berlin* **140**, 1013–1021.
7. Baker RF (1996) Abundance and species composition of zooplankton of the Nelson River estuary in relation to depth, tide, and location, July, 1995 - year 1. Unpublished report prepared by North/South Consultants Inc., Winnipeg, for Manitoba Hydro, Winnipeg, MB. viii + 65 p.

8. Bakker C (1980) On the distribution of ‘gonionemus vertens’ a. Agassiz (Hydrozoa, Limnomedusae), a new species in the eelgrass beds of Lake Grevelingen (S.W. Netherlands). *Aquatic Ecology* **14**, 186–195.

9. Bowman TE (1973) Pelagic amphipods of the genus Hyperia and closely related genera (Hyperiidea: Hyperiidae). *Smithsonian Contributions to Zoology* **136**, 1-76.

10. Bueckmann A (1969) Appendicularia. Fiches d'Identification du Zooplancton, 7: 1-9.

11. Cairns SD, Calder DR, Brinckmann-Voss A, Castro CB et al. (2002) Common and Scientific Names of Aquatic Invertebrates from the United States and Canada: Cnidaria and Ctenophora, Second Edition, 2002. American Fisheries Society Special Publication 28. xi + 115.

12. Choe N, Deibel D (2008) Temporal and vertical distributions of three appendicularian species (Tunicata) in Conception Bay, Newfoundland. J. *Plankton Research* **30**, 969-979.

13. Coad BW, Reist JD (2004) Annotated list of theArctic marine fishes of Canada. *Canadian Manuscript Report of Fisheries and Aquatic Sciences* **2674**, iv + 112p.

14. Coe WR (1956) Pelagic Nemertea: keys to Families and Genera. Fiches d'identification du zooplancton, 64. ICES: Copenhagen. 5 pp.

15. Cross W (1982) Under-ice biota at the Pond Inlet ice edge and adjacent fast ice areas during spring. *Arctic* **35**, 13–27.

16. Davis CC, Green JM (1974). Three monstilloids (Copepoda: Monstrilloida) from the Arctic. *Internationale Revue de gesamten Hydrobiologie und Hydrographie* **59**, 57–63.

17. Dunbar MJ (1942a) Marine macroplankton from the Canadian Eastern Arctic. I. Amphipoda and Schizopoda.—Canadian Journal of Research. Section D, *Zoological Sciences* **20**, 33–46.

18. Dunbar MJ (1942b) Marine macroplankton from the Canadian Eastern Arctic. II. Medusae, siphonophora, ctenophora, pteropoda, and chaetognatha. *Canadian Journal of Research* **20**, 71–77.

19. Dunbar MJ (1954) The amphipod Crustacea of Ungava Bay, Canadian eastern Arctic. *Journal of the Fisheries Research Board of Canada* **11**, 709–798.

20. Dunbar MJ (1963) Amphipoda. Sub-order Hyperiidea, Family Hyperiidae. *Fiches d'Identification du Zooplancton* **103**, 1–3.

21. Dunbar MJ (1964) Eupahusids and pelagic amphipods - Distribution in North Atlantic and Arctic waters. In: Serial atlas of the marine environment, Folio 6. American Geographical Society, New York, N.Y.

22. Evans MS, Grainger EH (1980) Zooplankton in a Canadian Arctic estuary, p. 199–210. In: V. S. Kennedy [Ed.]. Estuarine perspectives. Academic Press, New York, NY

23. Fenaux R, Bone Q, Deibel D (1998) Appendicularian distribution and zoogeography, p. 251-264. In q. Bone [ed.], The biology of pelagic tunicates. Oxford University Press.
24. Fishbase. Available at http://www.fishbase.org.
25. Forbes JR, Macdonald RW, Carmack EC, Iseki K, O'Brien MC (1992) Zooplankton retained by sequential sediment traps along the Beaufort Sea shelf break during winter. *Canadian Journal of Fisheries and Aquatic Science* **49**, 663–670.
26. Frost BW (1989) A taxonomy of the marine calanoid copepod genus Pseudocalanus. *Canadian Journal of Zoology* **67**, 525–551.
27. Gardner GA, Howell ET (1983) Zooplankton. Distribution across the shelf break of the southeast shoal of the Newfoundland Grand Banks in May 1981. *Canadian Manuscript Report of Fisheries and Aquatic Sciences* **1724**, iv + 61pp.
28. Gardiner K, Dick TA (2010) Arctic cephalopod distributions and their associated predators. *Polar Research* **29**, 209–227.
29. Grainger EH (1962) Zooplankton of Foxe Basin in the Canadian Arctic. *Journal of the Fisheries Research Board of Canada* **19**, 377–400.
30. Grainger EH (1965) Zooplankton from the Arctic Ocean and adjacent Canadian waters. *Journal of the Fisheries Research Board of Canada* **22**, 543–564.
31. Grainger EH (1975) Biological productivity of the Southern Beaufort Sea: the physical-chemical environment and the plankton. *Beaufort Sea Technical Report* **No. 12A**.
32. Grainger EH, Grohe C (1975) Zooplankton data from the Beaufort Sea, 1951 to 1975. *Canadian Fisheries Marine Service Technical Report* **591**, 1–54.
33. Harvey M, Starr M, Therriault JC, Saucier F, Gosselin M (2007) MERICA –Nord Program – monitoring and research in the Hudson Bay complex. Available at http://www.meds-sdmm.dfo-mpo.gc.ca/isdm-gdsi/azmp-pmza/docs/bulletin_5_07.pdf.
34. Hopcroft RR, Clarke C, Nelson RJ, Raskoff KA (2005) Zooplankton communities of the Arctic's Canada Basin: the contribution by smaller taxa. Polar Biology **28**, 198–206.
35. Hopky GE, Chiperzak DB, Lawrence MJ (1994a) NOGAP B2; Zooplankton, and larval and post larval fish data from Tuktoyaktuk Harbour, N.W.T., 1984 to 1987. *Canadian Data Report of Fisheries and Aquatic Sciences* **945**, v + 117 p.
36. Hopky GE, MJ Lawrence, Chiperzak DB (1994b) NOGAP B2; Zooplankton data from the Canadian Beaufort Sea shelf, 1987 and 1988. *Canadian Data Report of Fisheries and Aquatic Sciences* **912**, v + 219 p.
37. Hopky GE, Lawrence MJ, Chiperzak DB (1994c) NOGAP B2; Zooplankton data from the Canadian Beaufort Sea shelf, 1984 and 1985. *Canadian Data Report of Fisheries and Aquatic Sciences* **922**, iv + 164 p.

38. Hopky GE, Lawrence MJ, Chiperzak DB (1994d) NOGAP B2; Zooplankton data from the Canadian Beaufort Sea shelf, 1986. *Canadian Data Report of Fisheries and Aquatic Sciences* **923**, iv + 225 p.

39. Horner R, Murphy D (1985) Species composition and abundance of zooplankton in the nearshore Beaufort Sea in winter-spring. *Arctica* **38**, 201–209.

40. Jirkov IA (2001) [Polychaeta of the Arctic Ocean] Polikhety severnogo Ledovitogo Okeana. Moskva, Yanus-K, 1–632.

41. Kenchington ELR, Prena J, Gilkinson KD, Gordon DC Jr, MacIsaac K, Bourbonnais C, Schwinghamer PJ, Rowell TW, McKeown DL, Vass WP (2001) Effects of experimental otter trawling on the macrofauna of a sandy bottom ecosystem on the Grand Banks of Newfoundland. *Canadian Journal of Fisheries and Aquatic Sciences* **58**, 1043–1057.

42. Kosobokova KN, Hopcroft RR, Hirche H-J (2011) - Patterns of zooplankton diversity through the depths of the Arctic's central basins. *Marine Biodiversity* **41**, 29–50.

43. Kramp PL (1942) Siphonophora. The Godthaab Expedition 1928. *Meddeleser om Grønland* **80**, 3–24.

44. Kramp PL (1942) Medusae. The Godthaab Expedition 1928. *Meddeleser om Grønland* **81**, 1–168.

45. Maucheline J (1971) Euphausiacea adults. Fiches d'identification du zooplancton, 134. *ICES: Copenhagen*. 8 pp.

46. Levenstein RY (1981) [Some pecularities of the distribution of the family Polynoidae from the Canada basin of the Artic ocean]. Academy of Sciences of the USSR, Transactions of the P.P.Shirshov. *Institute of Oceanology* **115**, 26–36.

47. Markhaseva EL (2002) Phaennocalanus unispinosus (Copepoda, Calanoida, Phaennidae): new genus, and new species from the bathypelagial Arctic basin. *Sarsia* **87**, 312–318.

48. Mclaughlin *et al.* (2005) Common and scientific names of aquatic invertebrates from the United States and Canada: Crustaceans. American Fisheries Society, Special Publication 31, Bethesda, Maryland.

49. Mel'nikov IA, Kulikov AS (1983) The cryopelagic fauna of the central Arctic Basin [original in Russian]. Canadian Translation of Fisheries and Aquatric Sciences, 4910:1-92 (original publication 1980, Biol. Tsentral'nogo Arkticheskogo Basseina, p.97–111.

50. Muus BJ (1953a) Polychaeta, Families: Aphroditidae, Phyllodocidae and Alciopidae. Fiches d'identification du zooplancton, 52. *ICES: Copenhagen,* 6 pp.

51. Muus BJ (1953b) Polychaeta (contd.), Families: Tomopteridae and Typhloscolecidae. Fiches d'identification du zooplancton, 53. *ICES: Copenhagen,* 5 pp.

52. Ng ISY, Carr CM, Cottenie K (2009) Hierarchical zooplankton metacommunities: distinguishing between high and limiting dispersal mechanisms. *Hydrobiologia* **619**, 133–143.
53. North/South Consultants (2006) Potential dispersal of aquatic invasive species into Hudson Bay from ballast water from ships travelling from ports in Europe and North America. A report prepared for Fisheries and Ocean Canada File No. F2408-050083
54. Pepin P, Colbourne E, Maillet G (2011) Seasonal patterns in zooplankton community structure on the Newfoundland and Labrador Shelf. *Progress in Oceanography* **91**, 273–285.
55. Petryashov VV (2009) Order Mysidacea. In: Illustrated keys to the free-living invertebrates of Eurasian Arctic seas and adjacent deep waters. Vol. 1.: 65-83. Alaska Sea Grant College Program, University of Alaska Fairbanks.
56. Raskoff KA, Hopcroft RR, Kosobokova KN, Purcell JE, Youngbluth M (2010) Jellies under ice: ROV observations from the Arctic 2005 hidden ocean expedition. *Deep-Sea Research II*, 111-126.
57. Razouls C, de Bovée F, Kouwenberg J, Desreumaux N (2005-2015) - Diversity and Geographic Distribution of Marine Planktonic Copepods. Available at http://copepodes.obs-banyuls.fr/en.
58. Rogers GF (1981) Biological oceanography of Chesterfield Inlet, N.W.T., Summer 1978 – phytoplankton and zooplankton. Thesis, Univ. of Guelph, 185 pp.
59. Russell FS (1970) The Medusae of the British Isles. II. Pelagic Scyphozoa With a Supplement to the First Volume on Hydromedusae. Cambridge University Press, Cambridge. 284 pp.
60. Schuchert P (2001b) Survey of the family Corynidae (Cnidaria, Hydrozoa). *Revue Suisse de Zoologie* **108**, 739–878.
61. Schuchert P (2007) The European athecate hydroids and their medusae (Hydrozoa, Cnidaria): Filifera part 2. *Revue Suisse de Zoologie* **114**, 195–396.
62. Schuchert P (2010) The European athecate hydroids and their medusae (Hydrozoa, Cnidaria): Capitata Part 2. *Revue Suisse de Zoologie* **117**, 337–555.
63. Scott WB, Scott MG (1988) Atlantic Fishes of Canada. *Canadian Bulletin of Fisheries and Aquatic Sciences* **219**, 731p.
64. Shih CT (1977) A guide to the jellyfish of Canadian Atlantic waters.National Museum of Natural Scoences, Natural History Series, No.5, 90 pp.
65. Shih C-T, DR Laubitz (1978) Zooplankton distribution in the eastern Beaufort Sea and the Northwest Passage. *Astarte* **11**, 45–54.
66. Shih C-T, N Stallard (1982) Notes on two deep-water calanoids (*Aetideopsis rostrata* and *Neoscolecithrix farrani*) from Lancaster Sound. *Arctic* **35**, 56–60.

67. Shoemaker CR (1955) Amphipoda collected at the Arctic Laboratory, Office of Naval Research, Point Barrow, Alaska, by G. E. MacGinitie. *Smithsonian Miscellaneous Collections* **128**, 1–78.

68. Siferd TD, HE Welch, MA Bergmann, Curtis MF (1997) Seasonal distribution of sympagic amphipods near Chesterfield Inlet, N.W.T., Canada. *Polar Biology* **18**, 16–22.

69. Stephensen K (1933) The Godthaab Expedition 1928. Schizopoda. *Meddelelser om Grønland* **79**, 1–20.

70. Stübing D, Piepenburg D (1998) Occurrence of the benthic trachymedusa Ptychogastria polaris Allman, 1878 (Cnidaria: Hydrozoa) off Northeast Greenland and in the northern Barents Sea. *Polar Biology* **19**, 193–197.

71. Svavarsson J (1988) Bathyal and abyssal asollata (Crustacea, Isopoda) from the Norwegian, Greenland, and North Polar Seas. *Sarsia* **73**, 83–106.

72. Tattersall WM (1951) A review of the Mysidacea of the United States National Museum. *Bulletin of the United States National Museum* **201**, 1–292.

73. Tencati JR, Leung Y, Kobayashi H (1970) Taxonomic Guides to Arctic Zooplankton (I): Amphipods of the Central Arctic, Euphausiids of the Arctic Basin and Peripheral Seas. Technical rept. no. 2, 1959-1970, 61 pp.

74. Uchida T (1969) Medusae from the Arctic Ocean. *Publications of the Seto Marine Biological Laboratory* **17**, 285–287.

75. Walkusz W, Paulić JE, Kwaśniewski S, Williams WJ, Wong S, Papst MH (2010) Distribution, diversity and biomass of summer zooplankton from the coastal Canadian Beaufort Sea. *Polar Biology* **33**: 321–335.

76. Willey A (1920) Report on the marine Copepoda collected during the Canadian Arctic Expedition. Rep. *Canadian Arctic Expedition 1913-18* **7**, 1–46.

77. Willey A (1931) Biological and oceanographic conditions in Hudson Bay 4. Hudson Bay Copepod plankton. *Contributions Canadian Biology and Fisheries, N.S.* **6**, 483–493.

78. Woods SM, Smiley BD (1987) Arctic Data Compilation and Appraisal. Volume 9. Beaufort Sea: Biological Oceanography - Bacteria, Plankton, Epontic Community. Can. Data Rep. Hydrogr. *Ocean Science* **5**, 394 p.

79. WoRMS Editorial Board (2015). World Register of Marine Species. Accessed 2015-01-29. Available at http://www.marinespecies.org at VLIZ.

80. Zrum L (1999) Abundance and species composition of zooplankton in the Nelson River Estuary: baseline monitoring program 1998 – Year III. Unpublished report prepared by North/South Consultants Inc., Winnipeg, for Manitoba Hydro, Winnipeg, MB. viii + 71 p.

81. Zrum L (2000) Abundance and species composition of zooplankton in the Nelson River Estuary: baseline monitoring program 1999 – Year IV. Unpublished report

prepared by North/South Consultants Inc., Winnipeg, for Manitoba Hydro, Winnipeg, MB. ix+ 74 p

Appendix S3. Bioinformatic details

Specific commands for preparation and analysis of the paired-end reads:

1) **Trim:** forward and reverse sense fastq files for each sample were trimmed using trimmomatic 0.30 with the following parameters: (TrimmomaticPE, -phred33, ILLUMINACLIP:"$ADAPTERFILE":3:30:6, LEADING:20, TRAILING:20, SLIDINGWINDOW:20:20, MINLEN:200 2).
2) **Merge**: cleaned reads were then merged using FLASH v1.2.11 with the following options: (-t 1 -z -O -m 30 -M 280). Only merged reads were used in the following analyses.
3) **Separate reads for each primer separately:** since merged sample files contained amplicons for four pairs of primers (two COI pairs and two 18s pairs), we split them each by amplicon using a Python script (split_amplicons_one_file.py). The script looks for degenerate primers at the beginning and end of each sequences and, in the case of positive identification for both primers, puts the sequence in the appropriate file.
4) **Chimeras**: regroup all the samples of one amplicon in a single file and use usearch with the -uchime_denovo command and the default parameters and the following output options: (-uchimeout, -chimeras, -uchimealns).

Specific commands to assign reads at the species level using the *Barque pipeline*, an open source and freely available metabarcoding analysis pipeline (www.github.com/enormandeau/barque):

5) **Sequences reference:** To maintain the phylum information, all animal BIN databases have been downloaded separately prior to formating. Separate files prior to formating have been pooled in the file names (eg: Chordata.fas.gz). Python script (format_bold_database.py) has been used to format each databases. Briefly, we retained only COI sequences that contained well formatted genus and species names. We also provided a file containing a list of unwanted species containing the names of all insect species. The formatted databases contain sequences whose names are in the phylum_Genus_species format. We then concatenated the formatted BOLD database into a single Fasta file and created a usearch compatible database usearch -makeudb_usearch command.
6) **Multiple hits: t**he presence of multiple hits has been assessed for some of the sequences with usearch using the usearch_local command with the following options: (-id 0.95, -maxaccepts 10, -maxrejects 50, -strand both -blast6out, -

top_hits_only, -query_cov 0.5). We then ran the find_multiple_hits.sh script to parse the results and output a text file listing the cases of multiple hits and their frequency. The new unwanted species were added to the previous list of unwanted insect species, formatted and prepared the database again to remove them and used usearch with the usearch_local command with the following options to find the top hits only: (-id 0.9, -maxaccepts 6, -maxrejects 50, -strand both, -blast6out, -top_hit_only, -query_cov 0.5).

7) **Results** were summarized into phylum, genus and species count tables using the Python script 07_summarize_results.py with the following parameters: (min_similarity 97, min_length 300, min_coverage 1).

Specific commands to further explore missing species with the bold database using Operational taxonomic units (OTU) clustering:

5) **Read dereplication:** vsearch (--derep_fulllength --strand plus --sizeout --fasta_width 0)
4) **Operational taxonomic units (OTU) clustering:** OTUs were created using swarm 2.2.0 (-d 1, -f, -l,-w).
5) **Filtration:** OTUs represented by a single read (singletons) were excluded.
6) **Assignment:** the number of OTUs found with a ≤85% and ≤97% identity to a sequence in the BOLD database as well as the number of unique species retrived with at least 97% identity were found using vsearch (--usearch_global --qmask none --dbmask none --id 0.6 --blast6out --dbmatched --maxaccepts 20 --maxrejects 20 --maxhits 1 --query_cov 0.5 --fasta_width 0).
7) **Results** were summarized, for each phylum, site, and primer pair.

Table S2. Library preparation and bioinformatic pipeline details. Library details including the concentration (Conc., ng/µL) of the purified final PCR products measured by PicoGreen (ng/µL) and the final PCR volume mixed (µL) for each library (i.e. SRA accession number: SRX accession). Count of sequences for each data analysis step in the bioinfomatic pipeline (see Appendix S2) including the number of raw reads for the forward and reverse (i.e. count from the de-multiplexing), the reads remaining following trimming, merging (paired end runs), COI sequences within the expected amplicon length and without ambiguous nucleotides, without chimeras, and the sequences remaining with successful BLAST results (≥97% identity) and only for aquatic metazoan (removing insects, birds, algea and bacteria). Note that few non-aquatic species remained, but they were deleted before subsequent analyses. See the sampling site locations in Figure 1.

Port	Sample	SRA accession number	Sampling location	Sampling site	Library				Bioinformatic pipeline							
					COI1 (mlCOIintF-jgHCO2198)		COI2 (LCO1490-ill_C_R)		Number of raw reads for the forward side	Number of raw reads for the reverse side	Trimmed reads forward	Trimmed reads Reverse	Number of raw reads following merging	COI Sequences remaining with expected length and without ambiguous nucleotides	Sequences remaining without chimeras	Sequences with successful blast
					Conc.	Vol.	Conc.	Vol.								
Churchill	CH-11_S10	SRR5658897	surface	P2C	77.71	1.93	5.86	11.05	304083	304083	275237	275237	274260	117984	117984	2045
Churchill	CH-14_S34	SRR5658898	surface	P2B	72.33	2.07	6.40	10.11	324054	324054	291625	291625	290668	94641	94641	22011
Churchill	CH-17_S58	SRR5658903	surface	P8	75.07	2.00	13.93	4.65	281668	281668	253724	253724	252779	111175	111175	1589
Churchill	CH-20_S11	SRR5658894	surface	P7	71.43	2.10	5.03	12.86	290629	290629	263324	263324	262464	112530	112530	1644
Churchill	CH-21_S43	SRR5658794	surface	P9	72.15	2.08	14.29	4.53	255974	255974	232905	232905	232043	95948	95948	767
Churchill	CH-24_S67	SRR5658789	surface	P4	79.40	1.89	35.33	1.83	286480	286480	258167	258167	257188	115048	115048	8099
Churchill	CH-29_S44	SRR5659067	surface	P6	55.18	2.72	43.56	1.49	257483	257483	229346	229346	228586	85635	85635	1180
Churchill	CH-4_S25	SRR5659036	surface	P3B	75.22	3.20	22.31	2.90	280420	280420	255099	255099	253927	110392	110392	1985
Churchill	CH-63_S63	SRR5659127	surface	P1B	61.28	2.45	54.59	1.19	335859	335859	304633	304633	303687	128700	128700	1495
Churchill	CH-64_S55	SRR5659126	surface	P1C	56.73	2.64	77.63	0.83	369668	369668	336056	336056	335086	149025	149025	1327
Churchill	CH-69_S15	SRR5658835	surface	P5	71.35	2.10	12.64	5.12	276458	276458	251307	251307	250380	112146	112146	5920
Churchill	CH-7_S49	SRR5659041	surface	P10	76.39	1.96	1.31	20.00	269867	269867	242787	242787	241646	97353	97353	7076
Churchill	CH-12_S18	SRR5658896	mid-depth	P2C	51.13	2.93	1.10	20.00	241806	241806	215478	215478	213984	73443	73443	1223

Churchill	CH-15_S42	SRR5658901	mid-depth	P2B	62.30	2.41	0.17	20.00	206153	206153	181575	181575	180571	53026	53026	2710
Churchill	CH-18_S66	SRR5658902	mid-depth	P8	91.63	1.64	-0.07	20.00	245891	245891	217150	217150	215849	76825	76825	4570
Churchill	CH-22_S51	SRR5658787	mid-depth	P9	80.27	1.87	52.70	1.23	263808	263808	223190	223190	222380	100793	100793	74820
Churchill	CH-25_S4	SRR5658790	mid-depth	P4	71.61	2.09	29.24	2.21	307367	307367	275756	275756	274721	128933	128933	7622
Churchill	CH-28_S36	SRR5659068	mid-depth	P6	64.61	2.32	8.39	7.72	285379	285379	255398	255398	253995	110049	110049	9399
Churchill	CH-3_S17	SRR5659037	mid-depth	P3C	90.97	1.65	5.49	11.79	247825	247825	222966	222966	221826	85507	85507	5216
Churchill	CH-41_S27	SRR5658792	mid-depth	P7	76.40	1.96	6.89	9.40	238242	238242	202198	202198	200961	84956	84956	8909
Churchill	CH-5_S33	SRR5659043	mid-depth	P3B	76.61	2.85	32.56	1.99	255276	255276	230455	230455	229237	92707	92707	12974
Churchill	CH-62_S71	SRR5658737	mid-depth	P1B	59.22	2.53	29.13	2.22	302825	302825	275999	275999	275017	120872	120872	1275
Churchill	CH-65_S47	SRR5658843	mid-depth	P1C	68.73	2.18	14.28	4.53	257147	257147	231786	231786	230809	90495	90495	1197
Churchill	CH-68_S23	SRR5658838	mid-depth	P5	88.68	1.69	3.31	19.57	281421	281421	255179	255179	253969	102410	102410	1020
Churchill	CH-8_S57	SRR5659040	mid-depth	P10	96.13	1.56	25.99	2.49	255225	255225	222122	222122	221193	91489	91489	55880
Churchill	CH-13_S26	SRR5658899	deep	P2C	76.27	1.97	2.88	22.48	273525	273525	244825	244825	243692	90245	90245	11556
Churchill	CH-19_S8	SRR5658895	deep	P1B	51.94	2.89	4.61	14.05	280002	280002	250829	250829	249220	118388	118388	1436
Churchill	CH-23_S59	SRR5658788	deep	P9	61.68	2.43	3.02	21.46	290468	290468	263435	263435	262154	128607	128607	31255
Churchill	CH-26-1_S50	SRR5658900	deep	P2B	85.64	1.75	6.02	10.76	354669	354669	316253	316253	314635	128376	128376	69823
Churchill	CH-26-2_S20	SRR5658786	deep	P4	58.13	2.58	15.26	4.24	268010	268010	232266	232266	231238	96068	96068	43326
Churchill	CH-27_S28	SRR5659069	deep	P6	49.33	3.04	4.27	15.15	322194	322194	290609	290609	288711	139401	139401	2283
Churchill	CH-2_S9	SRR5659038	deep	P3C	89.30	1.68	12.94	5.01	265413	265413	236444	236444	234432	91531	91531	6683
Churchill	CH-42_S35	SRR5658793	deep	P7	42.50	3.53	-0.81	20.00	207513	207513	185573	185573	184249	51514	51514	3871
Churchill	CH-61_S3	SRR5658738	deep	P8	79.72	1.88	35.69	1.81	273227	273227	246720	246720	245912	99976	99976	2925
Churchill	CH-66_S39	SRR5658844	deep	P1C	67.71	2.22	4.88	13.26	260105	260105	234801	234801	233686	88421	88421	2442
Churchill	CH-6_S41	SRR5659042	deep	P3B	72.61	2.07	0.93	20.00	220835	220835	196403	196403	195110	64454	64454	9981
Churchill	CH-9_S65	SRR5659034	deep	P10	83.48	1.80	38.65	1.68	265596	265596	235010	235010	233724	98079	98079	46257
Churchill	CH-51_S72	SRR5659010	tide pool	P2A	57.73	2.60	52.21	1.24	270325	270325	235101	235101	233809	106108	106108	43369
Churchill	CH-52_S64	SRR5659011	tide pool	P2A	34.36	4.37	18.94	3.42	254212	254212	222889	222889	221744	90799	90799	27349
Churchill	CH-53_S56	SRR5659125	tide pool	P2A	62.73	2.39	20.80	3.11	294228	294228	255763	255763	254218	114780	114780	43346
Churchill	CH-54_S48	SRR5659124	tide pool	P2A	54.74	2.74	31.23	2.07	251524	251524	218595	218595	217271	90518	90518	23285
Churchill	CH-55_S40	SRR5658742	tide pool	P3A	50.97	2.94	17.05	3.80	259916	259916	226694	226694	225371	93146	93146	16422
Churchill	CH-56_S32	SRR5658741	tide pool	P3A	24.36	6.16	24.89	2.60	242107	242107	211763	211763	210701	88403	88403	50841
Churchill	CH-57_S24	SRR5658740	tide pool	P3A	29.91	5.02	1.18	20.00	211122	211122	186085	186085	185138	65226	65226	11644
Churchill	CH-58_S16	SRR5658739	tide pool	P3A	35.89	4.18	5.71	11.35	280065	280065	248902	248902	248026	110727	110727	19047

Churchill	CH-T76_S73	SRR5659009	tide pool	P1A	22.92	6.54	9.06	7.15	254859	254859	223607	223607	222868	101021	101021	5880
Churchill	CH-T80_S75	SRR5659007	tide pool	P1A	41.72	3.60	9.48	6.83	270627	270627	246870	246870	246204	113428	113428	1828
Churchill	CH-30_S62	SRR5658842	S20		60.53	2.48	37.25	1.74	316157	316157	281915	281915	280876	127631	127631	6690
Churchill	CH-31_S52	SRR5659066	S20		60.89	2.46	66.09	0.98	289459	289459	256737	256737	255914	124154	124154	10954
Churchill	CH-32_S60	SRR5659073	S20		64.34	2.33	30.07	2.15	294391	294391	264132	264132	263194	115087	115087	1877
Churchill	CH-33_S68	SRR5659072	S20		59.15	2.54	23.39	2.77	329335	329335	295439	295439	294236	119763	119763	1712
Churchill	CH-34_S5	SRR5659071	S20		36.60	4.10	35.43	1.83	266203	266203	239225	239225	238267	100788	100788	5486
Churchill	CH-35_S13	SRR5659070	S20		61.93	2.42	18.87	3.43	287440	287440	257201	257201	256317	101609	101609	1032
Churchill	CH-36_S29	SRR5659074	S20		60.51	2.48	47.79	1.35	163484	163484	118469	118469	117026	50396	50396	559
Churchill	CH-37_S37	SRR5658942	S20		70.01	2.14	58.15	1.11	271497	271497	241571	241571	240711	106118	106118	1565
Churchill	CH-38_S45	SRR5658943	S20		62.05	2.42	63.55	1.02	275186	275186	246189	246189	245304	94619	94619	1356
Churchill	CH-39_S53	SRR5658940	S20		74.35	2.02	55.06	1.18	304096	304096	274851	274851	273916	124608	124608	1861
Churchill	CH-40_S61	SRR5658941	S20		65.32	2.30	35.78	1.81	313520	313520	283189	283189	282254	125572	125572	2317
Churchill	CH-43_S69	SRR5658938	S20		58.07	2.58	46.35	1.40	296365	296365	265609	265609	264643	113080	113080	8688
Churchill	CH-44_S6	SRR5658939	S20		65.14	2.30	63.22	1.02	280096	280096	248653	248653	247714	101669	101669	1480
Churchill	CH-45_S14	SRR5658936	S20		66.25	2.26	35.68	1.81	321896	321896	289431	289431	288381	118989	118989	1604
Churchill	CH-46_S22	SRR5658937	S20		65.98	2.27	58.51	1.11	306955	306955	274497	274497	273454	114808	114808	2174
Churchill	CH-47_S30	SRR5658944	S20		57.79	2.60	33.65	1.92	286115	286115	257932	257932	257042	101014	101014	1791
Churchill	CH-48_S38	SRR5658945	S20		71.91	2.09	54.82	1.18	301879	301879	274344	274344	273472	115307	115307	1932
Churchill	CH-49_S46	SRR5658840	S20		63.84	2.35	36.47	1.78	262865	262865	233175	233175	232329	83916	83916	3037
Churchill	CH-50_S54	SRR5658839	S20		63.05	2.38	46.00	1.41	302460	302460	272847	272847	271879	119087	119087	1830
Churchill	CH-70_S70	SRR5658841	S20		56.98	2.63	38.23	1.69	296604	296604	267705	267705	266781	115535	115535	1691
Churchill	CH-T1_S2	SRR5659033	field control		-1.57	20.00	-0.91	20.00	10778	10778	7319	7319	7253	319	319	68
Churchill	CH-T2_S19	SRR5658791	field control		-1.82	20.00	-1.85	20.00	16438	16438	13618	13618	13602	350	350	9
Churchill	CH-T3_S12	SRR5658785	field control		-2.11	20.00	-2.33	20.00	4982	4982	3086	3086	3069	189	189	8
Churchill	CH-T4_S21	SRR5659075	field control		-1.98	20.00	-1.90	20.00	6414	6414	4507	4507	4496	257	257	11
Churchill	CH-T5_S7	SRR5658836	field control		-0.49	20.00	-1.33	20.00	66076	66076	56691	56691	56539	5923	5923	163
Churchill	CH-negA_S77	SRR5658797	extraction control		-1.56	20.00	-1.45	20.00	3496	3496	1386	1386	1377	48	48	6
Churchill	CH-negB_S78	SRR5658798	extraction control		-1.53	20.00	-2.24	20.00	1030	1030	217	217	215	40	40	2

Churchill	CH-negC_S79	SRR5658803	extraction control		-2.09	20.00	-0.72	20.00	2499	2499	607	607	600	69	69	12
Churchill	CH-negD_S80	SRR5658804	extraction control		-1.91	20.00	-2.25	20.00	488	488	129	129	124	55	55	2
Iqaluit	Iq-101_S67	SRR5659005	surface	P10	50.26	3.98	59.00	5.08	252040	252040	224417	224417	223884	83096	83096	1741
Iqaluit	Iq-112_S60	SRR5659110	surface	P1B	40.90	4.89	41.77	7.18	224744	224744	200101	200101	199649	78456	78456	85
Iqaluit	Iq-109_S72	SRR5659109	surface	P1C	46.85	4.27	43.37	6.92	157558	157558	140997	140997	140611	56204	56204	30
Iqaluit	Iq-106_S55	SRR5659102	surface	P2B	57.17	3.50	69.65	4.31	195743	195743	176091	176091	175676	78324	78324	928
Iqaluit	Iq-103_S8	SRR5659002	surface	P2C	45.36	4.41	52.21	5.75	160458	160458	143739	143739	143415	62816	62816	11
Iqaluit	Iq-140_S64	SRR5658784	surface	P3B	69.28	2.89	85.16	3.52	222426	222426	188990	188990	184864	70281	70281	7259
Iqaluit	Iq-143_S45	SRR5659025	surface	P3C	59.14	3.38	46.58	6.44	194700	194700	172040	172040	171469	60093	60093	421
Iqaluit	Iq-122_S41	SRR5659022	surface	P4	65.41	3.06	54.09	5.55	222326	222326	197071	197071	196512	62165	62165	191
Iqaluit	Iq-125_S23	SRR5658988	surface	P5	45.11	4.43	14.33	20.94	195855	195855	174917	174917	174299	47588	47588	341
Iqaluit	Iq-119_S62	SRR5658987	surface	P6	47.01	4.25	51.16	5.86	201715	201715	180333	180333	179513	79041	79041	168
Iqaluit	Iq-115_S59	SRR5658982	surface	P7	40.27	4.97	55.81	5.38	226991	226991	202491	202491	202145	79035	79035	687
Iqaluit	Iq-128_S28	SRR5659028	surface	P8	56.00	3.57	57.72	5.20	173578	173578	156210	156210	155788	69512	69512	22
Iqaluit	Iq-131_S53	SRR5659031	surface	P9	54.06	3.70	24.90	12.05	213189	213189	190672	190672	190181	69162	69162	2891
Iqaluit	Iq-102_S35	SRR5659004	mid-depth	P10	31.32	6.39	23.10	12.99	184840	184840	160550	160550	160105	54854	54854	1639
Iqaluit	Iq-113_S17	SRR5659111	mid-depth	P1B	60.18	3.32	68.28	4.39	230499	230499	204857	204857	204355	70555	70555	249
Iqaluit	Iq-110_S25	SRR5659106	mid-depth	P1C	64.29	3.11	81.78	3.67	209699	209699	187067	187067	186687	76911	76911	4157
Iqaluit	Iq-107_S26	SRR5659103	mid-depth	P2B	21.91	9.13	52.49	5.72	251386	251386	228422	228422	228066	131165	131165	1854
Iqaluit	Iq-104_S61	SRR5659104	mid-depth	P2C	66.56	3.00	62.64	4.79	225147	225147	200353	200353	199858	95848	95848	2473
Iqaluit	Iq-141_S57	SRR5658783	mid-depth	P3B	41.96	4.77	29.58	10.14	172687	172687	155541	155541	155130	72004	72004	485
Iqaluit	Iq-138_S24	SRR5659026	mid-depth	P3C	44.15	4.53	41.19	7.28	159793	159793	140656	140656	140287	48020	48020	188
Iqaluit	Iq-123_S54	SRR5659023	mid-depth	P4	57.68	3.47	50.34	5.96	230072	230072	206212	206212	205651	76817	76817	3236
Iqaluit	Iq-126_S12	SRR5658991	mid-depth	P5	40.34	4.96	36.22	8.28	219181	219181	189646	189646	189243	60145	60145	1073
Iqaluit	Iq-120_S3	SRR5658986	mid-depth	P6	52.54	3.81	60.45	4.96	238848	238848	212932	212932	212396	89410	89410	2745
Iqaluit	Iq-116_S65	SRR5658985	mid-depth	P7	53.94	3.71	36.90	8.13	230554	230554	204818	204818	204268	73360	73360	535
Iqaluit	Iq-129_S20	SRR5659029	mid-depth	P8	51.34	3.90	52.67	5.70	191758	191758	165394	165394	165010	52994	52994	681
Iqaluit	Iq-132_S69	SRR5658780	mid-depth	P9	51.47	3.89	50.08	5.99	166006	166006	148888	148888	148580	64286	64286	1342
Iqaluit	Iq-100_S85	SRR5659003	deep	P10	54.50	3.67	41.59	7.21	134167	134167	116274	116274	115883	51441	51441	3626
Iqaluit	Iq-114_S22	SRR5658983	deep	P1B	56.23	3.56	49.19	6.10	225276	225276	197713	197713	197166	66656	66656	422
Iqaluit	Iq-111_S6	SRR5659107	deep	P1C	47.92	4.17	32.15	9.33	214056	214056	187093	187093	186670	62668	62668	184
Iqaluit	Iq-108_S2	SRR5659108	deep	P2B	28.99	6.90	18.71	16.03	220824	220824	193034	193034	192532	58318	58318	1785

Iqaluit	Iq-105_S70	SRR5659105	deep	P2C	47.02	4.25	39.89	7.52	186276	186276	168545	168545	168088	76114	76114	290
Iqaluit	Iq-142_S4	SRR5658776	deep	P3B	41.15	4.86	67.36	4.45	243824	243824	213855	213855	213450	78897	78897	1671
Iqaluit	Iq-139_S39	SRR5659027	deep	P3C	47.13	4.24	55.28	5.43	144864	144864	131059	131059	130576	62942	62942	1033
Iqaluit	Iq-124_S63	SRR5659024	deep	P4	57.32	3.49	38.20	7.85	184808	184808	162412	162412	161934	72117	72117	9922
Iqaluit	Iq-127_S37	SRR5658990	deep	P5	68.97	2.90	32.59	9.21	148121	148121	130030	130030	129420	56361	56361	19
Iqaluit	Iq-121_S44	SRR5658989	deep	P6	54.80	3.65	74.15	4.05	177794	177794	155212	155212	154613	55825	55825	18016
Iqaluit	Iq-118_S11	SRR5658984	deep	P7	24.21	8.26	17.24	17.40	214428	214428	185291	185291	184614	50332	50332	27
Iqaluit	Iq-130_S27	SRR5659030	deep	P8	52.45	3.81	60.79	4.94	176268	176268	147163	147163	146748	61223	61223	41
Iqaluit	Iq-133_S32	SRR5658779	deep	P9	53.64	3.73	49.06	6.11	183341	183341	163821	163821	163478	60494	60494	329
Iqaluit	Iq-134_S52	SRR5658978	tide pool	P2A 1	43.43	4.61	81.80	3.67	250981	250981	215507	215507	214822	81188	81188	1502
Iqaluit	Iq-135_S15	SRR5658981	tide pool	P2A 2	51.55	3.88	59.89	5.01	235965	235965	207430	207430	206860	81372	81372	866
Iqaluit	Iq-136_S9	SRR5658980	tide pool	P2A 3	55.50	3.60	75.20	3.99	223808	223808	198735	198735	198146	80065	80065	3780
Iqaluit	Iq-137_S19	SRR5658973	tide pool	P2A-4	46.44	4.31	50.39	5.95	207396	207396	182475	182475	182113	68567	68567	670
Iqaluit	Iq-144_S18	SRR5658972	tide pool	P3A 1	48.19	4.15	49.96	6.00	231667	231667	203373	203373	202899	77937	77937	2490
Iqaluit	Iq-145_S13	SRR5658968	tide pool	P3A 2	41.68	4.80	56.16	5.34	177493	177493	150996	150996	150538	60568	60568	323
Iqaluit	Iq-146_S30	SRR5658969	tide pool	P3A 3	53.34	3.75	36.70	8.17	162433	162433	139784	139784	139107	50944	50944	814
Iqaluit	Iq-147_S50	SRR5658970	tide pool	P3A 4	46.95	4.26	46.88	6.40	251498	251498	222378	222378	221654	79222	79222	1472
Iqaluit	Iq-168_S92	SRR5658971	tide pool	P9A 1	28.58	7.00	19.03	15.76	149788	149788	125786	125786	125401	29514	29514	418
Iqaluit	Iq-169_S40	SRR5658964	tide pool	P9A 2	44.86	4.46	54.77	5.48	170103	170103	148818	148818	148445	53777	53777	122
Iqaluit	Iq-170_S14	SRR5658965	tide pool	P9A 3	45.86	4.36	35.47	8.46	228590	228590	201881	201881	201320	74786	74786	898
Iqaluit	Iq-171_S43	SRR5658966	tide pool	P9A 4	46.46	4.30	46.73	6.42	139641	139641	125058	125058	124725	55035	55035	714
Iqaluit	Iq-148_S33	SRR5658775	S20		25.77	7.76	20.10	14.93	150373	150373	134284	134284	134027	51972	51972	3029
Iqaluit	Iq-149_S38	SRR5658774	S20		43.48	4.60	44.67	6.72	147882	147882	134450	134450	134021	53909	53909	100
Iqaluit	Iq-150_S34	SRR5658773	S20		15.55	12.86	6.20	20.00	127989	127989	114601	114601	114237	29617	29617	12
Iqaluit	Iq-151_S42	SRR5658772	S20		39.53	5.06	50.69	5.92	204940	204940	181078	181078	180719	64575	64575	507
Iqaluit	Iq-152_S5	SRR5658781	S20		48.04	4.16	64.01	4.69	239878	239878	212161	212161	211726	82478	82478	178
Iqaluit	Iq-153_S10	SRR5658888	S20		50.88	3.93	50.48	5.94	241089	241089	214193	214193	213694	77123	77123	105
Iqaluit	Iq-154_S31	SRR5658889	S20		26.89	7.44	39.96	7.51	228105	228105	206273	206273	205677	74786	74786	80

Iqaluit	Iq-155_S48	SRR5658886	S20	41.90	4.77	31.74	9.45	154754	154754	136649	136649	136315	46011	46011	155
Iqaluit	Iq-156_S47	SRR5658887	S20	43.41	4.61	18.15	16.53	168259	168259	148600	148600	148203	43941	43941	76
Iqaluit	Iq-157_S71	SRR5658892	S20	37.66	5.31	14.93	20.09	181335	181335	161850	161850	161447	51188	51188	50
Iqaluit	Iq-158_S77	SRR5658893	S20	24.66	8.11	29.22	10.27	144767	144767	128158	128158	127738	43010	43010	85
Iqaluit	Iq-159_S66	SRR5658890	S20	34.15	5.86	50.14	5.98	214319	214319	188353	188353	187764	76693	76693	1231
Iqaluit	Iq-160_S21	SRR5658891	S20	51.68	3.87	47.66	6.29	218656	218656	194485	194485	193973	69922	69922	111
Iqaluit	Iq-161_S68	SRR5658880	S20	50.42	3.97	54.81	5.47	228295	228295	203174	203174	202730	85438	85438	123
Iqaluit	Iq-162_S29	SRR5658881	S20	43.96	4.55	52.68	5.69	102156	102156	67298	67298	66676	21649	21649	22
Iqaluit	Iq-163_S36	SRR5658975	S20	44.17	4.53	61.31	4.89	221426	221426	197383	197383	196860	65322	65322	119
Iqaluit	Iq-164_S58	SRR5658974	S20	41.21	4.85	50.23	5.97	258800	258800	230478	230478	229918	83002	83002	117
Iqaluit	Iq-165_S49	SRR5658977	S20	46.58	4.29	46.51	6.45	217246	217246	194100	194100	193611	79025	79025	177
Iqaluit	Iq-166_S81	SRR5658976	S20	51.11	3.91	39.71	7.55	192962	192962	170161	170161	169706	60544	60544	4040
Iqaluit	Iq-167_S73	SRR5658979	S20	46.10	4.34	43.75	6.86	234646	234646	202562	202562	201889	71245	71245	375
Iqaluit	Iq-T170_S7	SRR5658816	field control	5.62	20.00	2.89	20.00	25958	25958	21074	21074	21026	7823	7823	31
Iqaluit	Iq-T171_S46	SRR5658823	field control	4.11	20.00	1.64	20.00	1888	1888	607	607	602	195	195	2
Iqaluit	Iq-T172_S16	SRR5658824	field control	4.55	20.00	6.48	20.00	10185	10185	7416	7416	7409	668	668	1
Iqaluit	Iq-T173_S51	SRR5658920	field control	4.27	20.00	2.12	20.00	4889	4889	3191	3191	3179	683	683	15
Iqaluit	Iq-T174_S56	SRR5658919	field control	5.04	20.00	7.13	20.00	42576	42576	33812	33812	33776	4543	4543	59
Iqaluit	Iq-T175_S89	SRR5658922	field control	1.43	20.00	6.13	20.00	49880	49880	41068	41068	40985	959	959	25
Iqaluit	Iq-T176_S1	SRR5658921	field control	4.58	20.00	1.93	20.00	8460	8460	6200	6200	6183	1849	1849	0
Iqaluit	Iq-negA_S103	SRR5658801	extraction control	5.82	20.00	5.67	20.00	2624	2624	608	608	604	36	36	1
Iqaluit	Iq-negB_S76	SRR5658802	extraction control	5.73	20.00	5.44	20.00	1870	1870	1312	1312	1312	187	187	4
Iqaluit	Iq-negC_S80	SRR5658795	extraction control	5.57	20.00	5.46	20.00	2066	2066	1215	1215	1215	82	82	3
Iqaluit	Iq-negD_S84	SRR5658796	extraction control	5.65	20.00	5.89	20.00	1184	1184	445	445	444	128	128	1
Iqaluit	Iq-negE_S88	SRR5658904	extraction control	5.69	20.00	5.68	20.00	2844	2844	2014	2014	2009	155	155	9
Iqaluit	IqF-F1_S95	SRR5658967	F20	54.70	3.66	29.15	10.29	192490	192490	166238	166238	165347	59121	59121	2027
Iqaluit	IqF-F10_S96	SRR5659099	F20	45.14	4.43	37.20	8.06	192231	192231	166480	166480	165880	57871	57871	29718

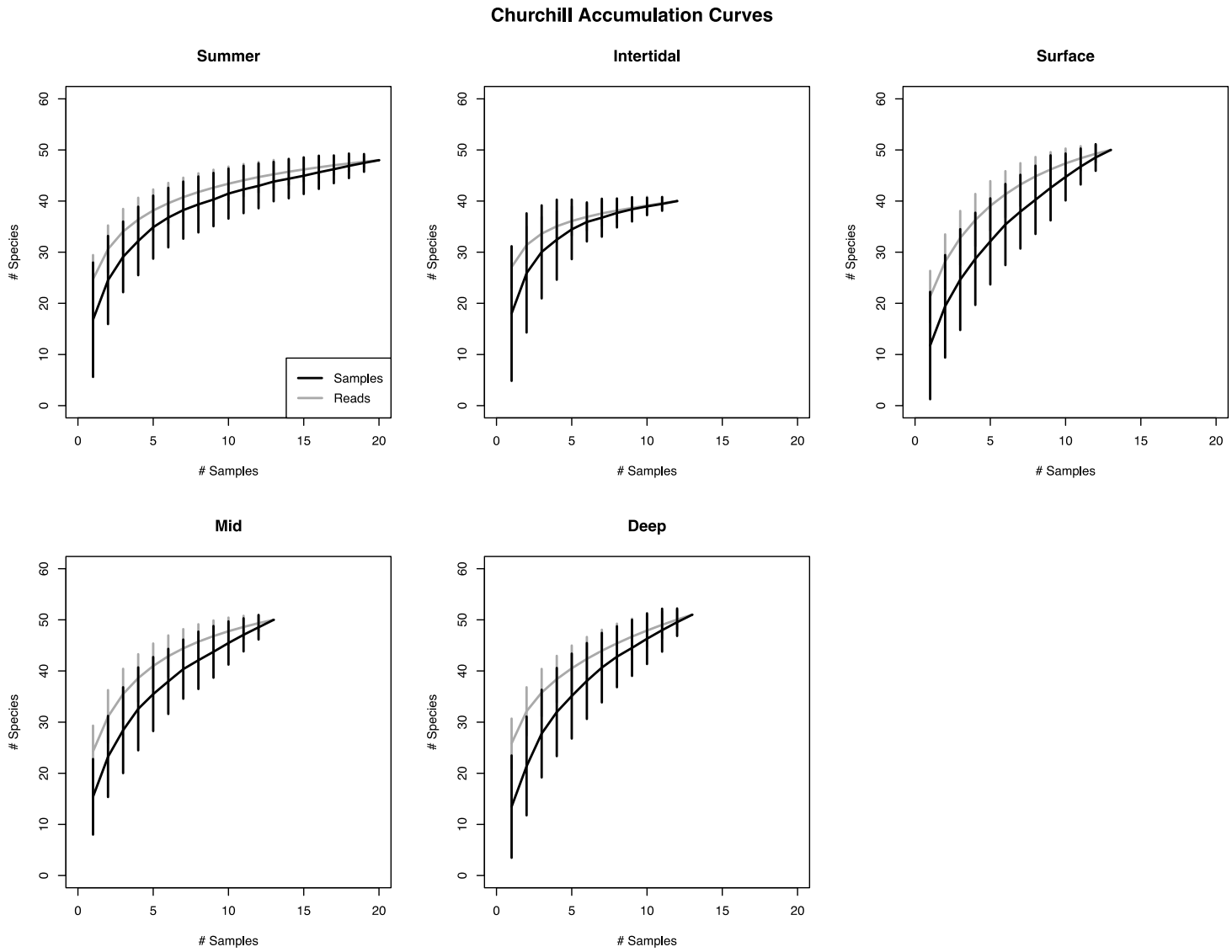
Iqaluit	IqF-F11_S99	SRR5659098	F20	39.97	5.00	23.13	12.97	68009	68009	60093	60093	60055	53427	53427	5705
Iqaluit	IqF-F12_S102	SRR5659089	F20	43.34	4.61	22.30	13.45	184325	184325	159916	159916	159188	45698	45698	844
Iqaluit	IqF-F13_S75	SRR5659088	F20	47.98	4.17	29.07	10.32	208630	208630	180086	180086	179143	46320	46320	526
Iqaluit	IqF-F14_S79	SRR5658821	F20	32.17	6.22	12.89	23.27	185048	185048	158156	158156	157446	30424	30424	1412
Iqaluit	IqF-F15_S83	SRR5658822	F20	49.48	4.04	20.75	14.46	189226	189226	163335	163335	162620	47309	47309	218
Iqaluit	IqF-F16_S87	SRR5658819	F20	30.33	6.59	13.86	21.65	161800	161800	139026	139026	138374	37423	37423	850
Iqaluit	IqF-F17_S91	SRR5658820	F20	21.40	9.35	14.15	21.20	161755	161755	135282	135282	134647	27991	27991	362
Iqaluit	IqF-F19_S94	SRR5658818	F20	37.97	5.27	23.29	12.88	157063	157063	133148	133148	132508	31824	31824	401
Iqaluit	IqF-F2_S98	SRR5658960	F20	42.63	4.69	22.58	13.29	163230	163230	136359	136359	135831	40977	40977	4799
Iqaluit	IqF-F20_S97	SRR5658815	F20	58.98	3.39	43.61	15.00	192105	192105	166470	166470	165759	56999	56999	33403
Iqaluit	IqF-F28_S100	SRR5658817	F20	54.71	3.66	24.49	15.00	190777	190777	164079	164079	163524	60096	60096	1179
Iqaluit	IqF-F3_S101	SRR5658961	F20	7.58	26.39	3.38	20.00	113748	113748	97960	97960	97396	17038	17038	172
Iqaluit	IqF-F4_S74	SRR5659097	F20	33.04	6.05	25.64	11.70	180462	180462	156350	156350	155623	60473	60473	4862
Iqaluit	IqF-F5_S78	SRR5659096	F20	51.04	3.92	21.83	13.74	181897	181897	156747	156747	156192	52576	52576	2939
Iqaluit	IqF-F6_S82	SRR5659095	F20	34.39	5.82	13.17	22.78	169616	169616	145468	145468	144899	40058	40058	734
Iqaluit	IqF-F7_S86	SRR5659094	F20	38.19	5.24	17.73	16.92	210834	210834	183220	183220	182386	58144	58144	781
Iqaluit	IqF-F8_S90	SRR5659101	F20	57.70	3.47	41.18	7.29	227529	227529	198144	198144	197382	61072	61072	802
Iqaluit	IqF-F9_S93	SRR5659100	F20	45.94	4.35	23.31	12.87	192197	192197	164600	164600	163831	40654	40654	681

Table S3. Reads with multiple species identified and actions taken in the pipeline based on the geographical species distributions.

Reads	Location	Phylum	Species	Action
26	Churchill	Mollusca	*Littorina saxatilis*	Keep
			Littorina compressa	Delete
			Littorina arcana	Delete
19	Churchill	Mollusca	*Littorina saxatilis*	Keep
			Littorina arcana	Delete
19	Churchill	Echinodermata	*Strongylocentrotus pallidus*	Strongylocentrotus sp.
22	Iqaluit		*Strongylocentrotus droebachiensis*	Strongylocentrotus sp.
14	Churchill	Porifera	*Baikalospongia bacillifera*	Delete
			Swartschewskia papyracea	Delete
			Lubomirskia baicalensis	Delete
			Ephydatia muelleri	Keep
			Baikalospongia recta	Delete
			Baikalospongia intermedia	Delete
14	Churchill	Chordata	*Coregonus nigripinnis*	Delete
			Coregonus artedi	Keep
			Coregonus hoyi	Delete
			Coregonus kiyi	Delete
			Coregonus zenithicus	Delete
11	Churchill	Mollusca	*Littorina saxatilis*	Keep
			Littorina compressa	Delete
4	Churchill	Anthropoda	*Hypogastrura viatica*	Delete
			Hypogastrura purpurescens	Delete
3	Churchill	Chordata	*Chen caerulescens*	Delete
			Chen rossii	Delete
1	Churchill	Porifera	*Swartschewskia papyracea*	Delete
			porifera Spongilla lacustris	Keep
1	Churchill	Porifera	*Ephydatia muelleri*	Keep
			Baikalospongia bacillifera	Delete
			Swartschewskia papyracea	Delete
			Lubomirskia baicalensis	Delete

			Baikalospongia recta	Delete
			Baikalospongia intermedia	Delete
1	Churchill	Porifera	*Ephydatia fluviatilis*	Ephydatia sp.
			Ephydatia muelleri	Ephydatia sp.
			Baikalospongia bacillifera	Delete
			Swartschewskia papyracea	Delete
			Lubomirskia baicalensis	Delete
			Baikalospongia recta	Delete
			Baikalospongia intermedia	Delete
1	Churchill	Arthropoda	*Daphnia pulex*	Daphnia pulex
			Daphnia pulicaria	Daphnia pulex
7	Iqaluit	Porifera	*Acanthorhabdus fragilis*	Porifera sp.
			Isodictya erinacea	Porifera sp.
			Iophon unicorne	Porifera sp.
5	Iqaluit	Chordata	*Lycodes seminudus*	Lycodes sp.
			Lycodes reticulatus	Lycodes sp.
			Lycodes lavalaei	Lycodes sp.
4	Iqaluit	Echinodermata	*Thyonidium drummondii*	Keep
			Thyone fusus	Keep
4	Iqaluit	Arthropoda	*Gammarus oceanicus*	Gammarus sp.
			Gammarus setosus	Gammarus sp.

455


Churchill Accumulation Curves
Summer
Intertidal
Surface
Mid
Deep
Species
Samples
Samples
Reads
0
5
10
15
20
0
10
20
30
40
50
60

456

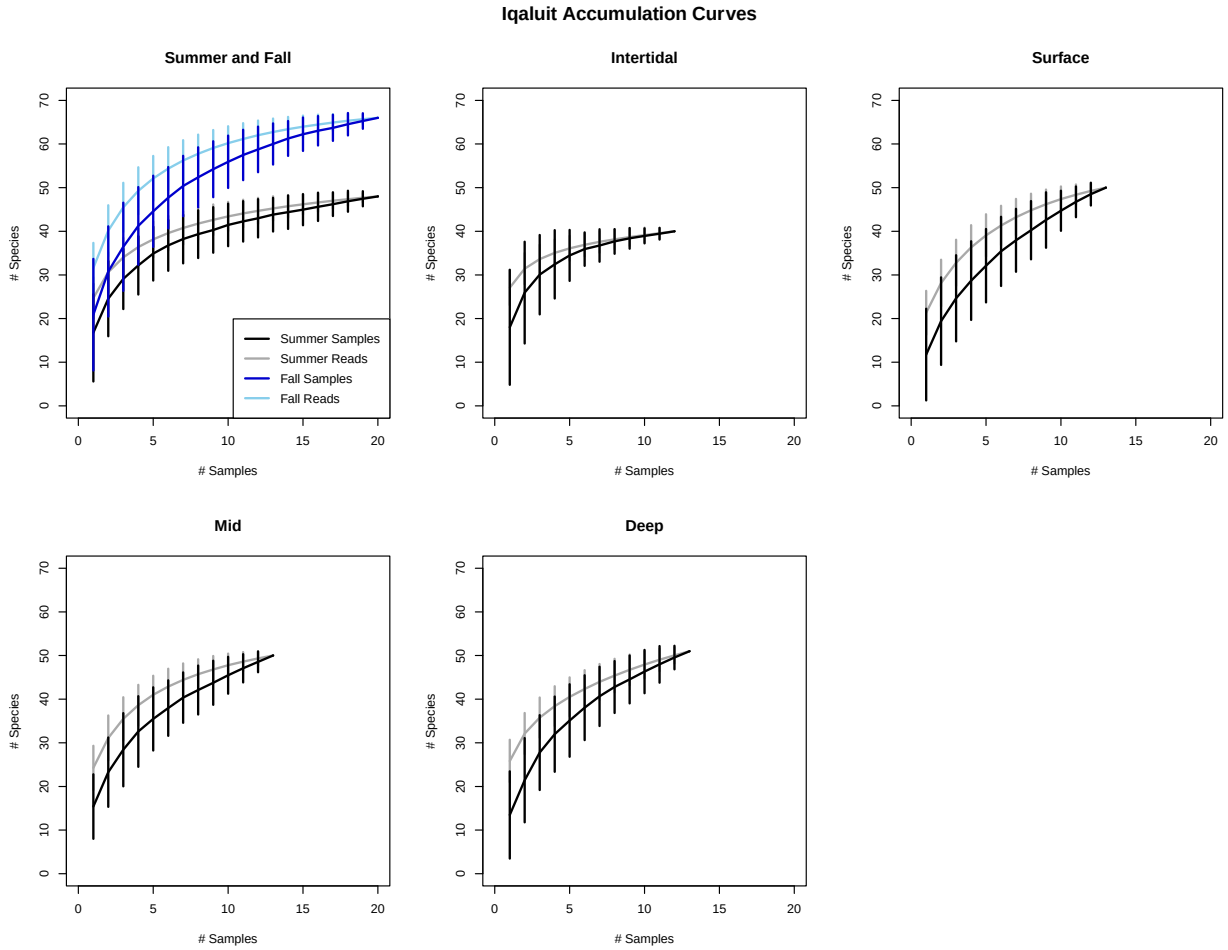
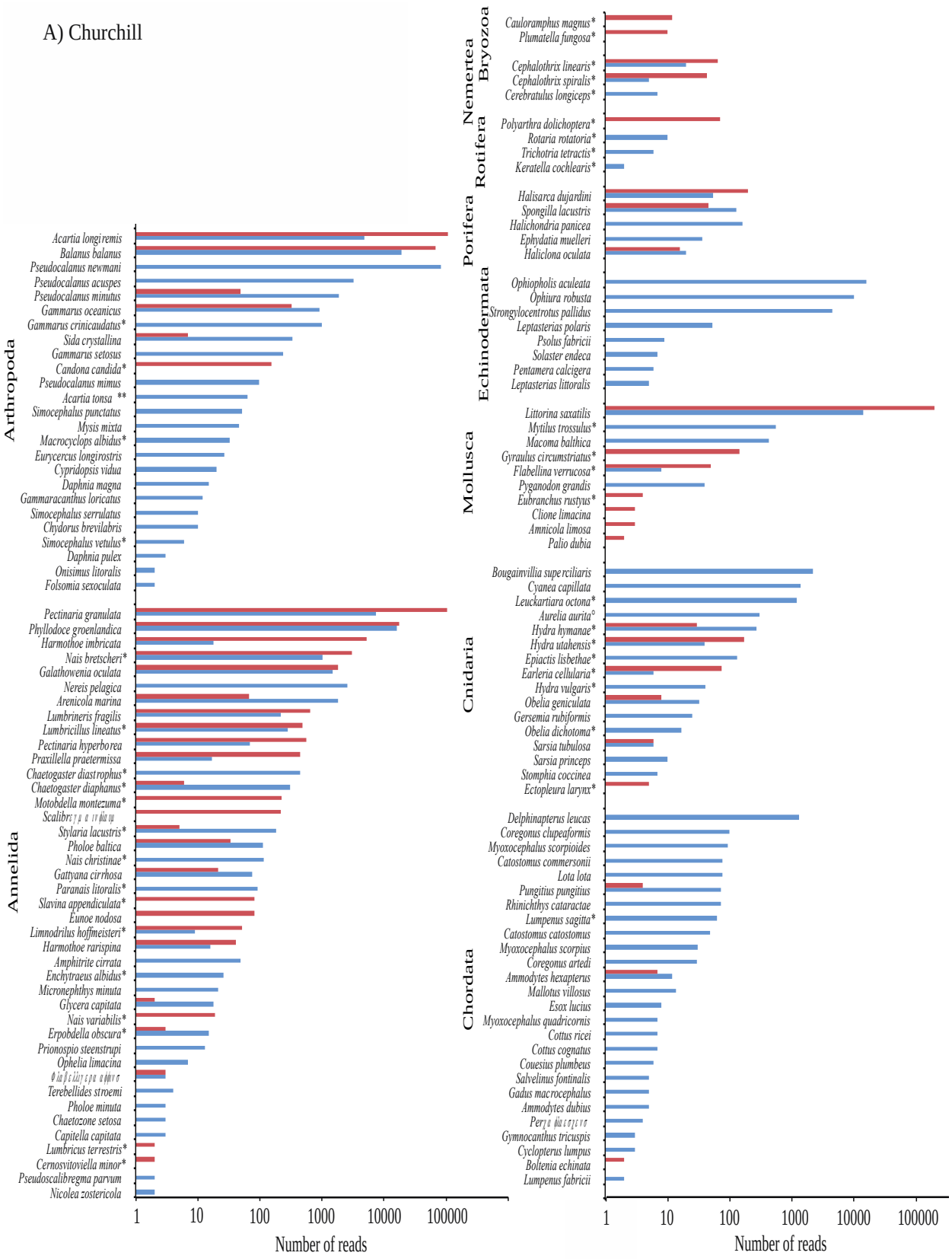


Figure S1. Rarefied species accumulation curves by read and sample numbers for each site for Churchill and Iqaluit. Read curves were plotted on the x-axis using the average number of reads per sample. Solid bold line denotes COI read rarefaction and light line denotes COI sample rarefaction. Errors bars represent 95% confidence intervals.


A) Churchill
Arthropoda
Acartia longiremis
Balanus balanus
Pseudocalanus newmani
Pseudocalanus acuspes
Pseudocalanus minutus
Gammarus oceanicus
Gammarus crinicaudatus*
Sida crystallina
Gammarus setosus
Candona candida*
Pseudocalanus mimus
Acartia tonsa **
Simocephalus punctatus
Mysis mixta
Macrocyclops albidus*
Eurycercus longirostris
Cypridopsis vidua
Daphnia magna
Gammaracanthus loricatus
Simocephalus serrulatus
Chydorus brevilabris
Simocephalus vetulus*
Daphnia pulex
Onisimus litoralis
Folsomia sexoculata
Annelida
Pectinaria granulata
Phyllodoce groenlandica
Harmothoe imbricata
Nais bretscheri*
Galathowenia oculata
Nereis pelagica
Arenicola marina
Lumbrineris fragilis
Lumbricillus lineatus*
Pectinaria hyperborea
Praxillella praetermissa
Chaetogaster diastrophus*
Chaetogaster diaphanus*
Motobdella montezuma*
Scalibregma inflatum
Stylaria lacustris*
Pholoe baltica
Nais christinae*
Gattyana cirrhosa
Paranais litoralis*
Slavina appendiculata*
Eunoe nodosa
Limnodrilus hoffmeisteri*
Harmothoe rarispina
Amphitrite cirrata
Enchytraeus albidus*
Micronephthys minuta
Glycera capitata
Nais variabilis*
Erpobdella obscura*
Prionospio steenstrupi
Ophelia limacina
Oligochaeta sp. unidentified
Terebellides stroemi
Pholoe minuta
Chaetozone setosa
Capitella capitata
Lumbricus terrestris*
Cernosvitoviella minor*
Pseudoscalibregma parvum
Nicolea zostericola
1
10
100
1000
10000
100000
Number of reads
Nemertea
Cauloramphus magnus*
Bryozoa
Plumatella fungosa*
Cephalothrix linearis*
Cephalothrix spiralis*
Cerebratulus longiceps*
Rotifera
Polyarthra dolichoptera*
Rotaria rotatoria*
Trichotria tetractis*
Keratella cochlearis*
Porifera
Halisarca dujardini
Spongilla lacustris
Halichondria panicea
Ephydatia muelleri
Haliclona oculata
Echinodermata
Ophiopholis aculeata
Ophiura robusta
Strongylocentrotus pallidus
Leptasterias polaris
Psolus fabricii
Solaster endeca
Pentamera calcigera
Leptasterias littoralis
Mollusca
Littorina saxatilis
Mytilus trossulus*
Macoma balthica
Gyraulus circumstriatus*
Flabellina verrucosa*
Pyganodon grandis
Eubranchus rustyus*
Clione limacina
Amnicola limosa
Palio dubia
Cnidaria
Bougainvillia superciliaris
Cyanea capillata
Leuckartiara octona*
Aurelia aurita°
Hydra hymanae*
Hydra utahensis*
Epiactis lisbethae*
Earleria cellularia*
Hydra vulgaris*
Obelia geniculata
Gersemia rubiformis
Obelia dichotoma*
Sarsia tubulosa
Sarsia princeps
Stomphia coccinea
Ectopleura larynx*
Chordata
Delphinapterus leucas
Coregonus clupeaformis
Myoxocephalus scorpioides
Catostomus commersonii
Lota lota
Pungitius pungitius
Rhinichthys cataractae
Lumpenus sagitta*
Catostomus catostomus
Myoxocephalus scorpius
Coregonus artedi
Ammodytes hexapterus
Mallotus villosus
Esox lucius
Myoxocephalus quadricornis
Cottus ricei
Cottus cognatus
Couesius plumbeus
Salvelinus fontinalis
Gadus macrocephalus
Ammodytes dubius
Perca flavescens
Gymnocanthus tricuspis
Cyclopterus lumpus
Boltenia echinata
Lumpenus fabricii
1
10
100
1000
10000
100000
Number of reads

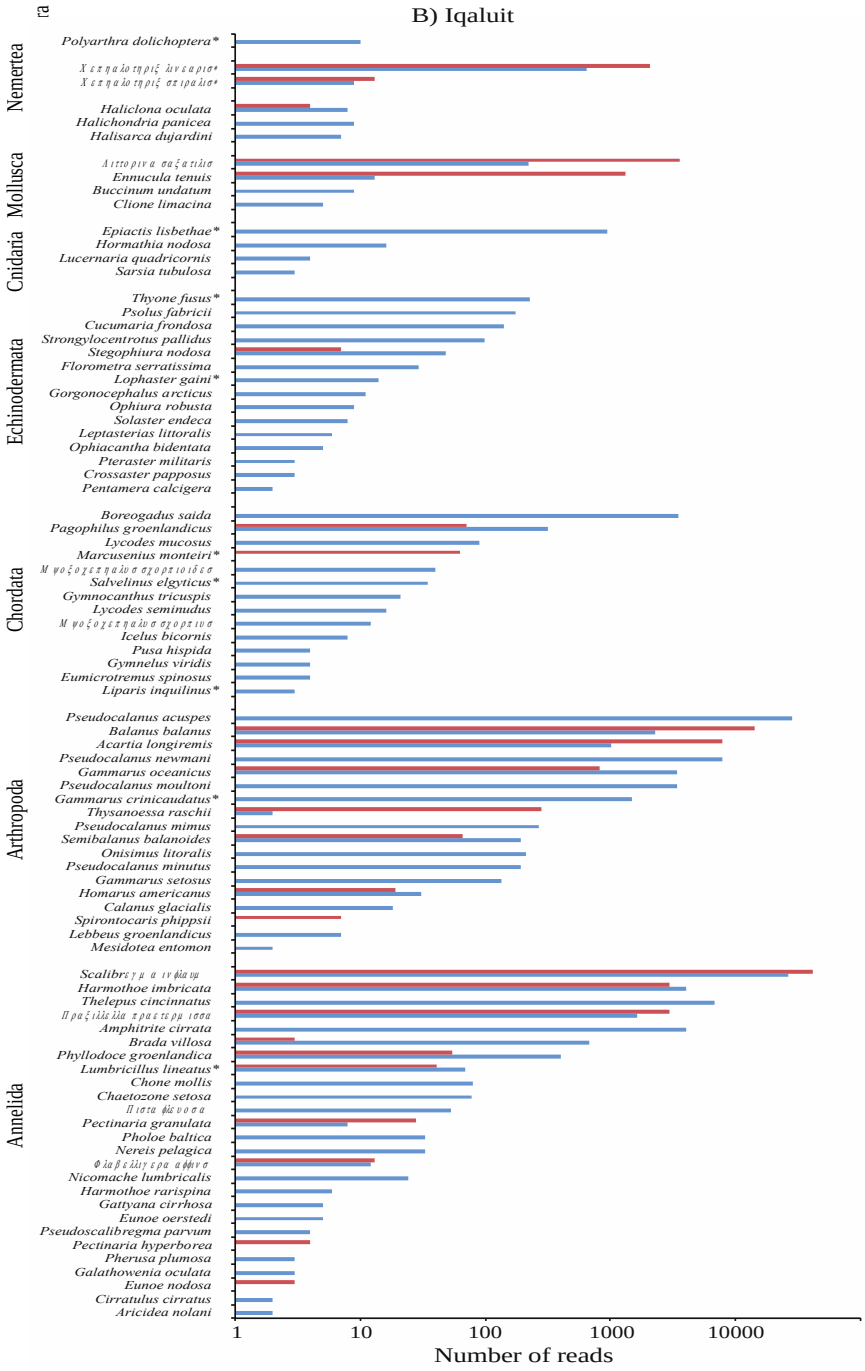


Figure S2. Species list and their known status within previous Canadian Arctic reported for each primers and port separately. The number of reads for each species detected from the eDNA (COI1 in blue and COI2 in red) collected in (A) Churchill and (B) Iqaluit in 2015. * indicate that this species was not previously detected in the Canadian Arctic, ** is a potential invader and ° only a single record of *Aurelia aurita* known from Canadian Arctic based on sequence data in BOLD from a partial specimen (from Churchill) that could not be morphologically identified to species level. Further research is needed to verify presence of this species in the Canadian Arctic.

Received: 14 April 2019 | Revised: 30 August 2019 | Accepted: 5 September 2019

DOI: 10.1002/edn3.35

ORIGINAL ARTICLE

Environmental DNA
Open Access
Dedicated to the study and use of environmental DNA for basic and applied sciences

WILEY

Comparing eDNA metabarcoding and species collection for documenting Arctic metazoan biodiversity

Noémie Leduc[1] | Anaïs Lacoursière-Roussel[2] | Kimberly L. Howland[3] | Philippe Archambault[1] | Maelle Sevellec[1,3] | Eric Normandeau[1] | Antoine Dispas[4] | Gesche Winkler[4] | Christopher W. McKindsey[5] | Nathalie Simard[5] | Louis Bernatchez[1]

[1]Département de Biologie, Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec, QC, Canada

[2]St. Andrews Biological Station (SABS), Fisheries and Oceans Canada, St. Andrews, NB, Canada

[3]Arctic Research Division Freshwater Institute, Fisheries and Oceans Canada, Winnipeg, MB, Canada

[4]Institut des Sciences de la mer (ISMER), Université du Québec à Rimouski, Rimouski, QC, Canada

[5]Institut Maurice-Lamontagne, Fisheries and Oceans Canada, Mont-Joli, QC, Canada

Correspondence
Noémie Leduc, Département de Biologie, Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec, QC G1V 0A6, Canada.
Email: noemie.leduc.2@ulaval.ca

Funding information
Nunavut Wildlife Management Board; Fisheries and Oceans Canada Aquatic Invasive Species Monitoring Program; Nunavik Marine Region Wildlife Board; POLAR Knowledge; ArcticNet

Abstract

Background: Arctic biodiversity has long been poorly documented and is now facing rapid transformations due to ongoing climate change and other impacts, including shipping activities. These changes are placing marine coastal invertebrate communities at greater risk, especially in sensitive areas such as commercial ports. Preserving biodiversity is a significant challenge, going far beyond the protection of charismatic species and involving suitable knowledge of the spatiotemporal organization of species. Therefore, knowledge of alpha, beta, and gamma biodiversity is of great importance to achieve this objective, particularly when partnered with new cost-effective approaches to monitor biodiversity.

Method and results: This study compares metabarcoding of COI mitochondrial and 18S rRNA genes from environmental DNA (eDNA) water samples with standard invertebrate species collection methods to document community patterns at multiple spatial scales. Water samples (250 ml) were collected at three different depths within three Canadian Arctic ports: Churchill, MB; Iqaluit, NU; and Deception Bay, QC. From these samples, 202 genera distributed across more than 15 phyla were detected using eDNA metabarcoding, of which only 9%–15% were also identified through species collection at the same sites. Significant differences in taxonomic richness and community composition were observed between eDNA and species collections at both local and regional scales. This study shows that eDNA dispersion in the Arctic Ocean reduces beta diversity in comparison with species collections while emphasizing the importance of pelagic life stages for eDNA detection.

Conclusion: The study also highlights the potential of eDNA metabarcoding to assess large-scale Arctic marine invertebrate diversity while emphasizing that eDNA and species collection should be considered as complementary tools to provide a more holistic picture of coastal marine invertebrate communities.

KEYWORDS
Arctic, beta diversity, biodiversity, eDNA, marine invertebrates, metabarcoding

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2019 The Authors. *Environmental DNA* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

The Arctic Ocean has been poorly surveyed and thus likely harbors a great undetected biodiversity (Archambault et al., 2010; Darnis et al., 2012). Recent estimates suggest that there are more than 4,000 species of invertebrates that inhabit the Arctic Ocean (Gradinger et al., 2010; Jørgensen, Archambault, Piepenburg, & Rice, 2016; Piepenburg et al., 2011) with greater than 90% being benthic organisms (CAFF International Secretariat, 2013). The general pattern of biodiversity decline with increasing latitude may not apply to marine invertebrates (Kendall, 1996), suggesting that a great diversity and many species await discovery (Archambault et al., 2010; Piepenburg et al., 2011). Archambault et al. (2010) showed that benthic infaunal diversity in the Canadian Arctic was almost similar than in Canadian Atlantic waters, even with three times less sampling effort. Previously, considered as the second most pristine oceans on earth (UNESCO, 2010), this ecosystem has experienced extensive environmental change since the 1950s (IPCC, 2018). In addition to warmer temperatures, increased acidification, and greater freshwater inputs (Arctic Climate Impact Assessment [ACIA], 2004), other activities such as marine shipping (ACIA, 2004; Chan et al.,) and the associated risk of introducing nonindigenous species (NIS) are increasing (Casas-Monroy et al., 2014; Chan, Bailey, Wiley, & MacIsaac, 2013; Goldsmit et al., 2018; Goldsmit, McKindsey, Archambault, & Howland, 2019). The number of invasive species has more than tripled since the beginning of the century in North America and in northern environments (Millennium Ecosystem Assessment, 2005; UNEP, 2006). Comprehensive baseline surveys and ongoing monitoring are thus essential in the Arctic, especially due to the large number of cryptic and cryptogenic species (Carlton, 1996; Goldsmit, Archambault, & Howland, 2014; Knowlton, 1993). However, gaining a better understanding of Arctic invertebrate community structure and how it may vary over time is challenging due to the heterogeneous distribution of species, uncertain taxonomy, and limitations due to sampling under ice cover (Jarosław, Mioduchowska, & Petković, 2016; Ministry of Environment, 2006).

The design of a robust monitoring approach to evaluate biodiversity changes, including species losses and processes that maintain species diversity over longer time frames, must take into account the spatial and temporal organization of biodiversity. Biodiversity can be measured using different taxonomy-based metrics and at various scales by evaluating alpha, beta, and gamma diversity. Alpha diversity represents the species assemblage of a relatively small area, termed "within-habitat diversity" (sensu MacArthur, 1965), and is the most commonly studied biodiversity scale. Beta diversity, often referred to as "turnover diversity," is the variation in species composition (i.e., species abundances and identities) among local species assemblages. It is the net outcome of regional biotic and abiotic processes, such as disturbance, the study of which may provide a mechanistic understanding of the processes that produce observed patterns and provide conservation-relevant insights on the maintenance of diversity over large spatial scales (McGill, Dornelas, Gotelli, & Magurran, 2015; Mori, Isbell, & Seidl, 2018; Socolar, Gilroy, Kunin, & Edwards, 2015). Lastly, gamma diversity refers to the species assemblage of large areas, for example, regional diversity (Socolar et al., 2015), and is expressed in the same units as alpha diversity (Laurila-Pant, Lehikoinen, Uusitalo, & Venesjärvi, 2015). Large-scale biodiversity monitoring is essential for understanding more extensive changes in coastal community composition, but this is logistically challenging and costly in remote areas such as the Arctic. Coastal metazoan collection methods are generally intrusive (e.g., trawling, grab sampling), selective, and frequently limited to the summer open water period and rely on some degree of subjectivity with respect to taxonomic expertise (Jones, 1992; Jørgensen et al., 2016).

Ten years after the pioneering study of Ficetola, Miaud, Pompanon, and Taberlet (2008), the environmental DNA (eDNA) approach offers major advantages over conventional monitoring methods and is perceived as a game-changer for ecological research (Creer et al., 2016). This approach involves the collection and detection of DNA that has been released by organisms into the surrounding environment through metabolic waste products, gametes, or decomposition (Hansen, Bekkevold, Clausen, & Nielsen, 2018; Taberlet, Bonin, Zinger, & Coissac, 2018). Analysis of eDNA with metabarcoding is a rapid method of biodiversity assessment that links taxonomy with high-throughput DNA sequencing (Ji et al., 2013) to provide a snapshot of local species composition without the need for sampling individual organisms. Recent studies in coastal marine ecosystems have demonstrated the feasibility of eDNA metabarcoding to document marine metazoan biodiversity in the Arctic (Grey et al., 2018; Lacoursière-Roussel et al., 2018). Despite limited knowledge of eDNA ecology (i.e., origin, fate, state, and transport; Barnes & Turner, 2016; Lacoursière-Roussel & Deiner, in press), eDNA is increasingly being incorporated within monitoring toolboxes for a large variety of aquatic organisms and ecosystems (Deiner et al., 2017; Roussel, Paillisson, Tréguier, & Petit, 2015).

However, like any sampling approach, eDNA metabarcoding also has its weaknesses which must be considered to avoid misinterpretation of results. Although the tool allows rapid assessment of biodiversity, database gaps hamper the use of eDNA as sequence assignments are highly dependent on their availability in public databases (Elbrecht, Vamos, Meissner, Aroviita, & Leese, 2017; Kwong, Srivathsan, & Meier, 2012). Organism detection is also restricted by the primers used and their respective biases (Elbrecht & Leese, 2015). Furthermore, unlike direct species collection, eDNA does not provide any physiological or health information for the detected organisms (Thomsen & Willerslev, 2015).

In the aquatic realm, while many studies have compared species composition measured by eDNA with conventional methods for fish (Thomsen et al., 2012; Yamamoto et al., 2017), few such comparative studies have been performed on invertebrates, and even less have considered the spatial scales of observation. Among marine invertebrate species, meroplankton (organisms having planktonic larval life stages) and holoplankton (organisms spending their entire life as plankton) represent key components of the food web and ecosystem stability (Gajbhiye, 2002; Marcus & Boero, 1998). A better

understanding of how complex planktonic life stages of invertebrates affect the origin and transport of eDNA in coastal environments is essential to develop genomics-based biodiversity indices to inform conservation plans.

The main objective of this study is to compare patterns of biodiversity at different spatial scales revealed by eDNA metabarcoding and conventional species collection within and among three ports in the Canadian Arctic Ocean. More specifically, gamma biodiversity (species richness between ports) was compared based on results from eDNA and conventional collecting methods, namely benthic trawl, Van Veen grab, cores, and plankton net tows. Secondly, alpha (species richness within ports) and beta (similarity of species between sites within ports) biodiversity indices were contrasted for results based on eDNA and species collections, to better understand how eDNA may inform species distributions and ecological processes such as dispersion and biotic heterogenization or homogenization. Finally, the life histories of organisms were considered to interpret how this basic biological parameter may affect eDNA detections from coastal invertebrates and contribute to discrepancies between eDNA detection and conventional species collections.

2 | MATERIALS AND METHODS

2.1 | Sample collection

Individual specimens from traditional sampling methods and eDNA were collected at 13 subtidal stations (≤20 m at low tide) in three commercial harbors of the Canadian Arctic in summer (Figure 1). Churchill was surveyed 11–14 August 2015, Iqaluit between 17–22 August 2015 and 24–26 July 2016, and Deception Bay between 19 and 27 August 2016. These three Arctic ports were selected because of their risk to potential changes in their local marine invertebrate communities due to climate change and the relatively high levels of shipping activity in each, which places them at greater risk for introduction of nonindigenous species (Chan et al., 2013; Chan et al., ; Goldsmit et al., 2019).

2.1.1 | Species collection

Throughout the paper, we use *specimens collected* and *species collection* to refer to the following collecting methods: benthic trawls, Van Veen grabs, sediment cores, and plankton tows. We use the term benthic communities to refer to organisms collected through benthic trawls, Van Veen grabs, and sediment cores, while we use the term zooplankton to refer to organisms collected using net tows. Benthic invertebrates living on the sea floor substrate (epifauna) were collected using a benthic trawl with a 500-µm-mesh net, while benthic invertebrates living in soft sea bottoms (infauna) were collected using a Van Veen grab (0.1 m^2 sample area; Deception Bay and Iqaluit) with the contents sieved on a 500-µm mesh. Zooplankton was collected using 0.5-m-diameter net tows: one vertical 80 µm and one oblique 250 µm. Zooplankton samples were taken at 10 of the 13 stations where eDNA was sampled, whereas benthic trawl and Van Veen grab samples were taken at all 13 stations. Trawling and oblique net tows were carried out for 3 min at a speed of 1–2 knots. Due to logistical constraints, Iqaluit Van Veen and trawl samples were collected in 2015 and 2016, respectively. Infauna samples in Churchill were collected by divers using corers (15 cm high × 10 cm diameter) from the same areas used by Goldsmit (2016). Since the sediment volume accumulated by these subtidal sediment cores was less than that of the Van Veen grab, the replicates of a given site for the sediment cores were combined together such that the final volume included for analyses was similar to the volume of site-specific Van Veen grab samples from the other ports. With the exception of common easily identifiable macroinvertebrates, which were enumerated, recorded, and released, all specimens were preserved in 95% ethanol and later identified by trained taxonomists to the lowest taxonomic level possible.

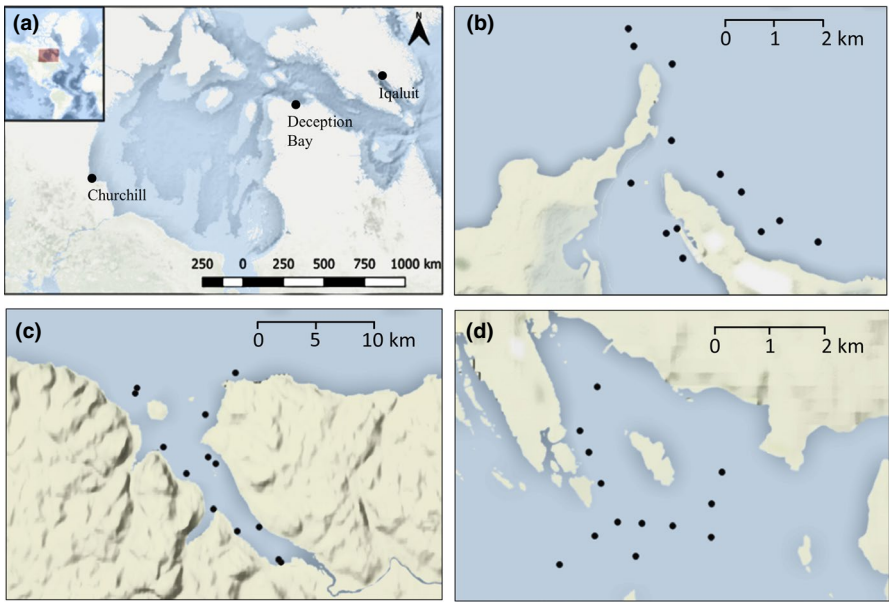


FIGURE 1 Geographic location of Churchill, Deception Bay, and Iqaluit harbors in the Canadian Arctic (a) and distribution of stations within Churchill (b), Deception Bay (c), and Iqaluit (d)

2.1.2 | Environmental DNA samples

A total of 117 water samples were collected and filtered following the methods outlined in Lacoursière-Roussel et al. (2018). A 250-ml water sample was taken at each of the three depths (surface, mid-depth, and deep water [i.e., 50 cm from the bottom]) for each station and port using 5-L Niskin bottles. The surface water was collected within the first meter, whereas mid-depth samples were collected at an average depth of 7.2 m (*SD* = 1.9), 6.8 m (*SD* = 2.8), and 9.8 m (*SD* = 3.5) for Churchill, Deception Bay, and Iqaluit, respectively, while deep-water samples were collected at an average depth of 12.7 m (*SD* = 2.7) and 15.5 m (*SD* = 4.6) for the same port, respectively. Each sample was filtered in the field using a 0.7-µm glass microfiber filter (Whatman GF/F, 25 mm) and syringes (BD 60 ml, Franklin Lakes, NJ, USA). Negative field controls were made by filtering 250 ml of autoclaved distilled water for every 10 collected samples. All filters were preserved in 2-ml microtubes containing 700 µl of Longmire's lysis/preservation buffer, kept at 4°C until the end of a sampling campaign, and then frozen at −20°C until extraction (at most 4 months). Risks of cross-contamination during the field sampling process were reduced by using a separate sterile kit for each sample. Sampling kits included bottles and a filter housing sterilized with a 10% bleach solution and new sterilized gloves, syringes, and tweezers sealed in a transparent plastic bag. Each sampling kit was exposed to UV light for 30 min following assembly.

2.2 | Metabarcoding

2.2.1 | Environmental DNA extraction, amplification, and sequencing

To avoid risk of laboratory cross-contamination, eDNA extraction, PCR preparation, and post-PCR steps were done in three separate rooms. All PCR manipulations were done in a decontaminated UV hood. All laboratory bench surfaces were cleaned with DNA AWAY®, and all laboratory tools were sterilized with a 10% bleach solution and exposed to UV light for 30 min before any manipulations were carried out. DNA was extracted from filters following a QIA shredder and phenol/chloroform protocol (Lacoursière-Roussel et al., 2018). Negative control extractions (950 µl distilled water) were done for each sample batch (i.e., one for every 23 samples) and were treated as normal samples for the remaining manipulations until sequencing. No positive controls were done in the context of this study since the efficiency of the selected primers used was previously tested on 104 zooplankton species and was validated on mock metazoan communities collected in Canadian ports by Zhang (2017). Furthermore, the primer sequences were also previously evaluated in silico with sequence databases for their ability to detect native and potential non-indigenous Arctic metazoans by Lacoursière-Roussel et al. (2018).

To maximize biodiversity detection and reduce the bias of eDNA dominance among species groups, two pairs of primers from two different genes (COI and 18S) were used. These have been shown to work well for detecting a wide variety of taxa including invertebrates and have reasonably comprehensive databases of reference sequences. Following Lacoursière-Roussel et al. (2018), we used the forward mlCOIintF (Leray et al., 2013) and reverse jgHCO2198 (Geller, Meyer, Parker, & Hawk, 2013) (hereafter called COI1) and the forward LCO1490 (Folmer, Black, Hoeh, Lutz, & Vrijenhoek, 1994) and reverse ill_C_R (Shokralla et al., 2015) (hereafter called COI2). Two additional universal 18S primer pairs were also used, the forward F-574 and reverse R-952 (Hadziavdic et al., 2014) (hereafter called 18S1) and the forward TAReuk454FWD1 and reverse TAReukREV3 (Stoeck et al., 2010) (hereafter called 18S2). Three PCR replicates were done for each sample of each primer set and were then pooled following amplification and purification (see Data S1 for more details). Sequencing was carried out using an Illumina MiSeq (Illumina) with a paired-end MiSeq Reagent Kit V3 (Illumina) at the Plateforme d'Analyses Génomiques (IBIS, Université Laval, Québec, Canada). Each port was analyzed on a separate run to ensure independence, but the samples within a port were pooled within a single Illumina MiSeq run to ensure the equality of sequencing depth among samples. Raw sequence reads were deposited in NCBI's Sequence Read Archive (SRA, http://www.ncbi.nlm.nih.gov/sra) under Bioprojects PRJNA388333 and PRJNA521343.

2.2.2 | Bioinformatics

Adaptor and primer sequences were removed and raw sequencing reads demultiplexed into individual samples files using the MiSeq Control software v2.3. Raw reads were analyzed using Barque version 1.5.1, an eDNA metabarcoding pipeline (www.github.com/enormandeau/barque). Forward and reverse sequences were trimmed and filtered using Trimmomatic v 0.30 with the following parameters: TrimmomaticPE, -phred33, LEADING: 20, TRAILING: 20, SLIDINGWINDOW: 20:20, and MINLEN: 200 (Bolger, Lohse, & Usadel, 2014). Pairs of reads were merged with FLASh v1.2.11 (Fast Length Adjustment of Short reads) with the following options: -t 1 -z -m 30 -M 280 (Magoč & Salzberg, 2011). The amplicons were split using their primer pairs (COI1, COI2, 18S1 and 18S2), and sequences that were either too short or too long were removed. Chimeric sequences were removed using VSEARCH v 2.5.1 (uchime_denovo command with default parameters) (Rognes, Flouri, Nichols, Quince, & Mahé, 2016). COI sequences were blasted on the BOLD database and 18S sequences against the SILVA database. Sequences from most terrestrial species (insects, human, birds, and mammals) and sequences that had no taxonomic match were also removed from the reference databases. Finally, following these steps, chordates others than tunicates (Table S1) were removed from the results since they were not targeted in this study and would therefore blur the analyses and subsequent interpretations regarding invertebrate communities. The Barque pipeline (https://github.com/enormandeau/barque) was then used to create operational taxonomic units (OTU). The OTUs were generated using VSEARCH 2.5.1 (id 0.97) (https://github.com/torognes/vsearch) using only reads present more than 20 times in the full dataset due to its meaningful size. For each station, sequences

collected at different depths and for all primers were pooled to obtain an overall representation of potential biodiversity.

2.3 | Data analysis

All analyses were performed at the genus level to facilitate comparisons between the approaches since only ~60% and 80% of the invertebrate taxa could be identified to species level with species collections and the eDNA approach, respectively. All analyses were done using R version 3.4.3 (R Core team, 2017) except for the SIMPER analyses which were done using PRIMER 6 and PERMANOVA+ (Clarke and Gorley, 2006).

In order to determine the effect of sampling effort on overall detected richness, genus-level rarefaction curves were created for each port and data collection type using the "specaccum" function in the R *vegan* package (Oksanen et al., 2016). Variation in taxonomic composition detected with eDNA and species collection within ports was depicted using a barplot generated in R from the raw relative abundance of genus taxonomy matrices assigned to a corresponding phylum. PERMANOVAs (number of permutations = 10,000) were performed using the *vegan* package to test the effect of port and sampling method on taxonomic composition, while nonmetric multidimensional scaling (nMDS) was used to visualize differences in taxonomic composition among ports and sampling methods.

Using an integrative approach based on the data at hand, alpha diversity indices (richness, Shannon diversity *H′*, and Pielou evenness *J*) were calculated using the R *vegan* package (Oksanen et al., 2016) following the Hellinger standardization. Variations in diversity indices between ports and sampling methods were evaluated using two-way ANOVAs followed by the Tukey honestly significant difference (Tukey HSD) tests. When standard ANOVA assumptions of normality were not met, PERMANOVAs were done based on Euclidean distances, thereby ensuring approximate multivariate normality (Clarke & Warwick, 2001), followed by pairwise comparisons using the "pairwise.adonis" function in R to evaluate variation in diversity due to sampling approaches among ports.

Beta diversity was estimated from the Sorensen distance using the "vegdist" function in the *vegan* package (Oksanen et al., 2016) computed based on presence–absence data. Geographic distance matrices between stations within ports were calculated using the "spDistsN1" function in the R *sp* package (Bivand, Pebesma, & Gomez-Rubio, 2008) for Deception Bay and Iqaluit, while distance between Churchill stations was determined using ArcGIS version 10.4 due to some peculiarities of the geographic layout of this port (this port has a large peninsula separating some sample stations; Figure 1b, and as *sp* simply calculates the straight-line distance between two points, the distances between stations on either side of this peninsula are underestimated using *sp*, whereas ArcGIS allows for calculation of the true distance by water). The dispersion of eDNA within ports was evaluated from correlations between beta diversity and spatial distance matrices using Mantel tests in the R *ade4* package (Dray & Dufour, 2007) except for Churchill for which the correlation was calculated using the "cor.test" function (method = Spearman) in the R stats package as ArcGIS does not provide a suitable distance matrix format for the Mantel test.

Finally, we investigated the probability of detecting different marine invertebrate taxa according to their life cycle, paying particular attention to those including pelagic stages (holoplankton and meroplankton) due to their potential presence in the water column. To contrast the proportion of species with an entirely pelagic (i.e., holoplankton) versus benthic–pelagic (i.e., meroplankton) life cycles, a barplot was constructed in R from a presence/absence data list with the lowest taxonomic resolution for each organism and the associated life cycle category. Variation in taxonomic composition among ports within each life history type (holoplankton vs. taxa with meroplanktonic life stages) was assessed using PERMANOVA using the *vegan* package. Similarity percentage analysis (SIMPER) in PRIMER 6 and PERMANOVA+ was used to determine which taxa contributed the most to explaining differences among groups.

3 | RESULTS

3.1 | Sequencing quality

A total of 478,046 aquatic metazoan reads were obtained in Churchill, 95,658 in Deception Bay, and 203,245 in Iqaluit (see Table S2 for further details on pipeline processes). The 18S markers generally generated more sequences than did COI markers, except for Iqaluit where the opposite trend was observed (Table 1). Genus-level taxonomic resolution provided a satisfactory description of biodiversity given that less than 20% were not assigned at this taxonomic level in all locations (Figure S1). Thus, a total of 2,682, 1,413, and 1,056 operational taxonomic units (OTUs) were identified at the genus level in the ports of Churchill, Deception Bay, and Iqaluit, respectively.

No amplification was observed on agarose gels for the negative PCR controls, but a small number of sequences were present in our laboratory and field negative controls (Table S3). Two correction factors were applied to ensure the reliability of the data and quality of the resulting analyses. First, the few sequences present in the laboratory negative controls were subtracted from the samples from the same extraction batch. These sequences represent 0.003%, 0.1%, and 0.06% of Churchill, Deception Bay, and Iqaluit total number of sequences, respectively. Second, for the negative field controls, a genus was removed if its abundance in all the field controls was greater than 2% of the total number of sequences for all field samples combined for that genus. This percentage threshold was established considering that the removal of genera with a contamination between 0% and 2% would have led to an erroneous representation of marine invertebrates detected by eDNA. Following application of correction factors for background contamination, 0.1% and 1.4% of all COI and 18S sequences, respectively, were removed (Table S4). An exception to applying correction was made in the case of 18S *Pseudocalanus* sequences for which 96% of all the field contamination occurred in only one field negative control. Given that *Pseudocalanus* in real samples represented nearly half of all 18S sequences and this genus is known to be a dominant part of the Arctic

TABLE 1 Summary of the numbers of reads, the proportion of species and genera present in the historic (i.e., previously described from) Arctic database, and the mean number of OTUs for the COI primer set and the 18S primer set that are assigned and nonassigned on BOLD and SILVA for each harbor

	Number of reads		Proportion of species known in Arctic (%)		Proportion of genera known in Arctic (%)		Mean no. of assigned OTUs (genus)		Mean no. of nonassigned OTUs (genus)	
Harbor	COI	18S	COI	18S	COI	18S	COI	18S	COI	18S
Churchill	52,749	425,297	52.3	18.7	61.7	45.9	633	708	39	100
Deception Bay	30,214	65,454	62.9	18.3	74.3	52.6	348	359	16	105
Iqaluit	125,104	78,141	69.4	15.4	77.6	46.3	238	291	4	92

Note: The list of described species in the Arctic was obtained by pooling various species databases (N = 1,054 species; K.L. Howland, P. Archambault, N. Simard and R. Young, unpublished data) and published information (Cusson, Archambault, & Aitken, 2007; Goldsmit et al., 2014; Link, Chaillou, Forest, Piepenburg, & Archambault, 2013; López et al., 2016; Olivier, San Martín, & Archambault, 2013; Piepenburg et al., 2011; Roy, Iken, & Archambault, 2015; Young, McCauley, Galetti, & Dirzo, 2016).

zooplankton community (Dispas, 2019), removing it would significantly bias the analyses. When read abundance of a given genus in field controls was lower than 2% of the total number of sequences for that genus, it was retained because contamination was considered low enough that it would not lead to false interpretations. In contrast, discarding those genera could bias analyses due to their high number of sequences in real samples.

3.2 | Arctic coastal gamma diversity

With the exception of benthos communities sampled using trawls, grabs, and cores, genera rarefaction curves of marine invertebrates were close to saturation for both zooplankton and eDNA (Figure S2). A total of 634 marine invertebrate genera from 23 phyla were identified when eDNA and species collection datasets were combined. Gamma richness was consistently higher for species collections methods (432 genera identified) than for eDNA (202 genera detected), and there was variation between sampling approaches among ports. eDNA gamma richness was higher for Churchill and Deception Bay but lower for Iqaluit, whereas the opposite pattern was observed for the gamma richness of communities detected with species collection (Figure 2a). Although a substantial collective number of organisms were detected, few genera were shared between eDNA and species collections (Churchill 15%, Deception Bay 15%, and Iqaluit 9%). Of the organisms found with both approaches, annelids accounted for almost half (42.7%), followed by arthropods and mollusks with 20.2% and 11.2%, respectively, of the common genera obtained within all ports (Figure 2b).

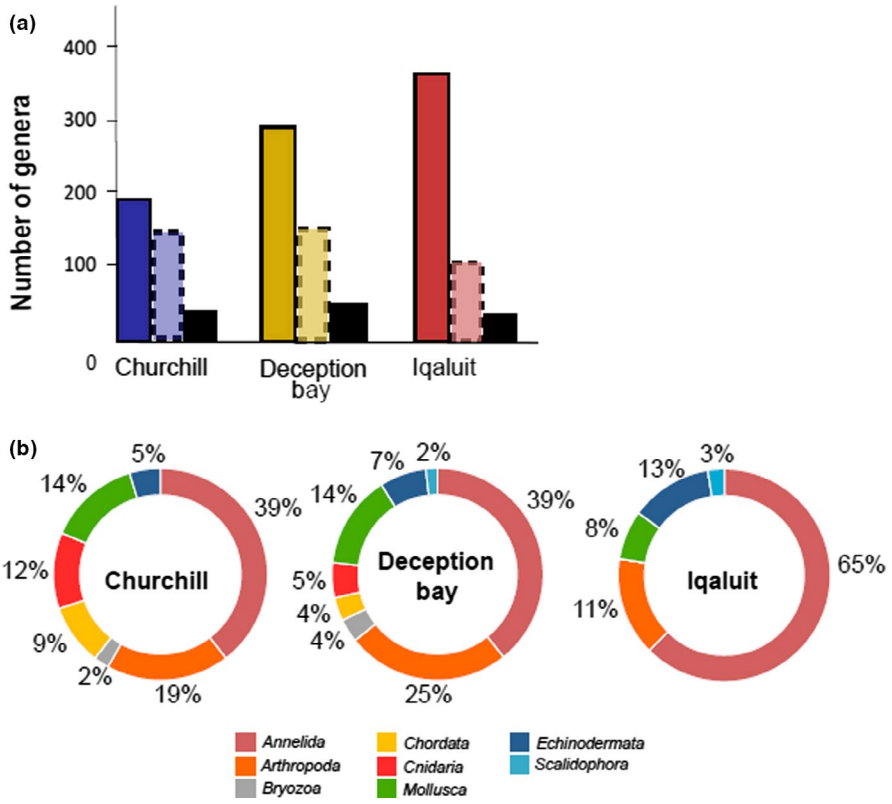


FIGURE 2 (a) Barplots of gamma richness (the total number of genera found) in Churchill (blue), Deception Bay (yellow), and Iqaluit (red). Darker bars represent species collection methods, whereas pale bars with dashed outlines represent eDNA and black bands represent the number of genera in common between the two collection methods. (b) Relative proportion of common genera identified by eDNA and species collection methods by phylum. Data represent pooled COI and 18S primer and traditional collection methods datasets for both (a) and (b)

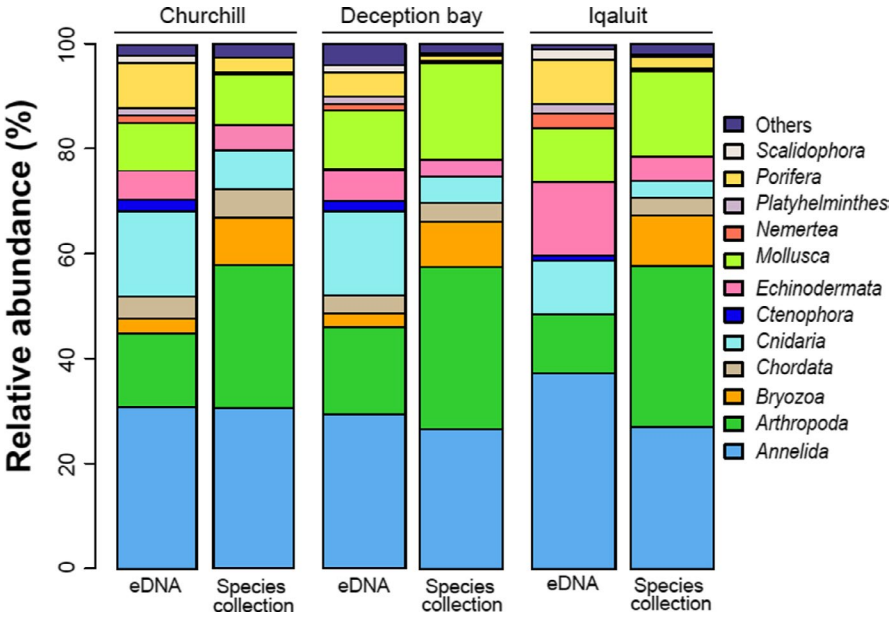


FIGURE 3 Marine invertebrate taxonomic composition at the phylum level for eDNA and species collection methods, respectively, for the ports of Churchill, Deception Bay, and Iqaluit ports. The COI and 18S datasets and benthic trawl, core, Van Veen grab, and net tow datasets are pooled for the eDNA and species collection barplots, respectively

The same phyla were generally present among the three ports, with *Annelida* and *Arthropoda* consistently being the most abundant phyla for both eDNA and species collections. However, the relative abundance of most taxa differed significantly between eDNA and species collections (PERMANOVA, $p < .001$; Table S5; Figure 3). Community composition of eDNA clearly differed among ports (PERMANOVA, $p < .001$; Table S5; Figure 4a) as did, although less clear visually, that for species collections (PERMANOVA, $p < .001$; Table S5; Figure 4b). Differences in community structure with eDNA versus species collection were mainly driven by *Annelid* and *Arthropod* genera (SIMPER analysis; 30% and 23%, respectively), followed by mollusks, echinoderms, cnidarians, and bryozoans (SIMPER analysis; 11%, 6%, 5%, and 4%, respectively). The remaining differences between eDNA and species collection community compositions may be partly driven by taxon-specific differences in detectability by these approaches. For example, some taxa such as *Brachiopoda*, *Foraminifera*, *Cephalorhyncha*, and *Chaetognatha* (grouped in the *Others* category with additional phyla of low relative abundance) were only found using species collection, while others such as *Bryozoa* were only rarely detected using eDNA. In contrast, taxa such as *Porifera*, *Nemertea*, *Cnidaria*, and *Echinodermata* were more frequently detected with higher read abundances in eDNA samples than in species collections.

3.3 | Arctic coastal alpha biodiversity

As for gamma diversity, alpha richness for eDNA samples was significantly higher in Churchill and Deception Bay than in Iqaluit (Tukey HSD, $p < .01$), with the number of genera per station ranging from 49 to 75 (mean = 63 ± 2) in Churchill, 45 to 93 (mean = 70 ± 4) in Deception Bay, and 34 to 53 (mean = 41 ± 2) in Iqaluit (Figure 5a). In contrast, Churchill had the lowest alpha richness for species collection samples (Tukey HSD, $p < .01$; Figure 5b) with only 8–58 genera per station (mean = 27 ± 3) as compared to 30–142 (mean = 78 ± 9) and 59–151 (mean = 100 ± 8) genera per station in Deception Bay and Iqaluit, respectively. Overall differences between sampling approaches varied between ports, with eDNA-based alpha richness being higher than species collection sample-based richness in Churchill (PERMANOVA, $p < .001$; Table S5), similar in Deception Bay (PERMANOVA, $p = .4$; Table S5), and lower in Iqaluit (PERMANOVA,

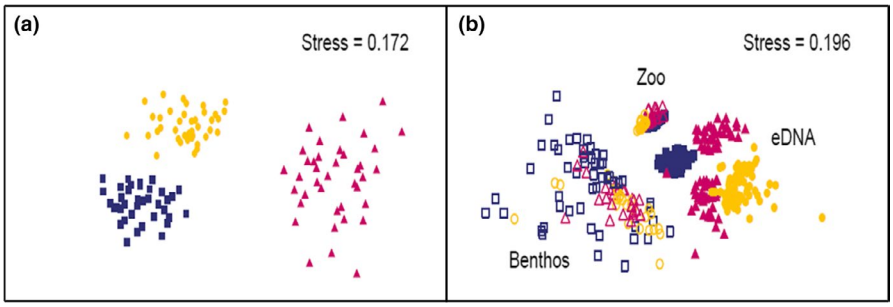


FIGURE 4 Variation in biodiversity (a) among ports based on eDNA and (b) among sampling methods within ports. Ordination of taxonomic composition (genera) calculated using the Sorensen index (incidence based) with each data point representing a sample. Blue squares represent Churchill, yellow circles Deception Bay, and magenta triangles Iqaluit. Filled and hollow symbols represent eDNA and species collection samples, respectively

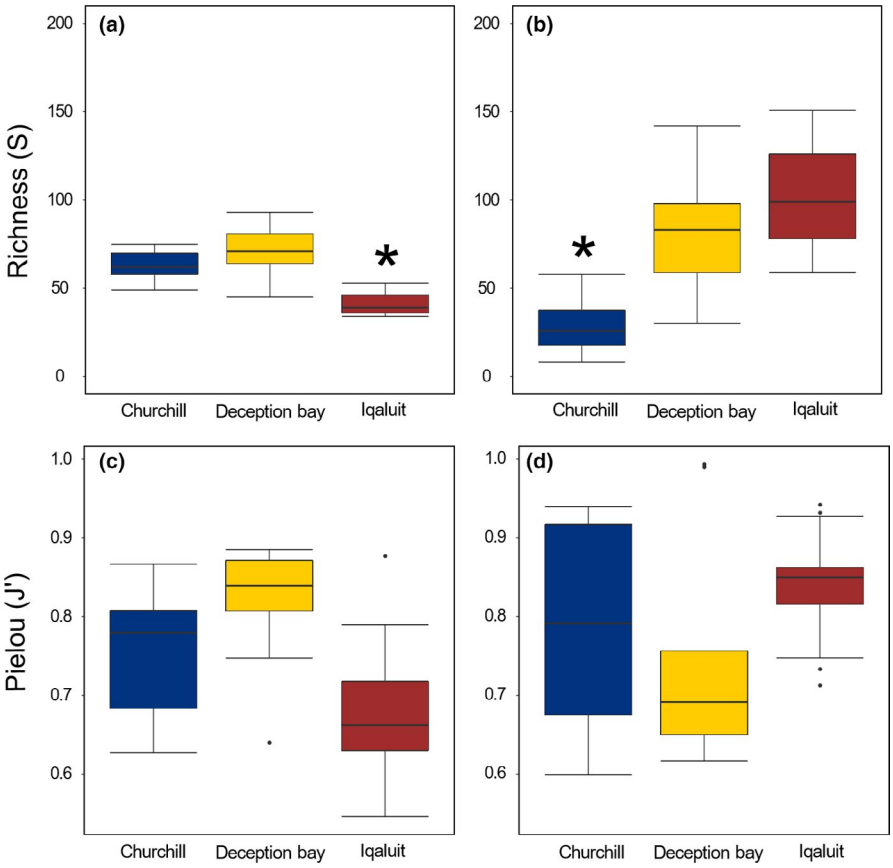


FIGURE 5 Boxplots of alpha diversity for genus-level richness and Pielou evenness index in Churchill, Deception Bay, and Iqaluit harbors for eDNA (a, c) and species collection (b, d). The COI and 18S datasets and benthic trawl, core, Van Veen grab, and net tow datasets are pooled for the eDNA and species collection boxplots, respectively. Significantly different richness are marked with an *

$p < .001$; Table S5). A similar pattern was observed for the Shannon biodiversity index (Figure S3).

Despite the contrasting alpha richness between sampling approaches within each port, the generally high values of Pielou's evenness indices revealed a pronounced taxonomic evenness with little indication of particular genera being overrepresented in communities detected by eDNA or species collection methods within the studied ecosystems (Table 2). Community evenness evaluated with eDNA was similar across ports except between Deception Bay and Iqaluit, where a lower or greater dominance by some taxa was observed in Iqaluit (PERMANOVA, $p < .05$; Table S5; Figure 5c). This is consistent with the SIMPER analyses where, for Iqaluit, 19 genera explained 90% of the similarity among stations in contrast to 30 and 42 genera for Churchill and Deception Bay, respectively. There were no differences in community evenness detected in species collections among the three ports (PERMANOVA, $p = .2$; Table S5; Figure 5d).

3.4 | Arctic coastal beta diversity

Community structure between stations within ports differed significantly for both eDNA and species collection but was greater for species collections than eDNA (Table 2). For eDNA, highest dissimilarity among stations was found in Iqaluit (0.37 ± 0.005), followed by Deception Bay (0.33 ± 0.005) and Churchill (0.31 ± 0.004), while the opposite trend was observed for species collections (Churchill: 0.84 ± 0.008; Deception Bay: 0.62 ± 0.01; Iqaluit: 0.58 ± 0.007).

TABLE 2 Summary of richness and alpha and beta biodiversity indices for eDNA and species collection of marine invertebrate communities on abundance data following Hellinger (Shannon and Pielou indices) and presence/absence (beta index) transformations, respectively. The COI and 18S datasets and benthic trawl, core, Van Veen grab, and net tow datasets are pooled for the eDNA and species collection datasets, respectively

Method	Harbor	Gamma richness ($S\gamma$)	Mean alpha richness ($S\alpha$) ± *SE*	Mean Pielou (*J*) ± *SE*	Mean Shannon (H') ± *SE*	Beta index ± *SE*
eDNA	Churchill	138	63 ± 2	0.75 ± 0.02	3.12 ± 0.1	0.31 ± 0.004
	Deception Bay	145	70 ± 4	0.82 ± 0.02	3.48 ± 0.1	0.33 ± 0.005
	Iqaluit	101	41 ± 2	0.67 ± 0.03	2.50 ± 0.1	0.37 ± 0.005
Species collection	Churchill	193	27 ± 3	0.79 ± 0.02	2.50 ± 0.1	0.84 ± 0.008
	Deception Bay	292	78 ± 9	0.75 ± 0.04	3.17 ± 0.1	0.62 ± 0.01
	Iqaluit	365	100 ± 8	0.84 ± 0.02	3.84 ± 0.1	0.58 ± 0.007

Positive correlations between beta diversity and geographic distance between stations were observed for most eDNA and species collections across all ports. Positive correlations between distance and eDNA beta diversity were significant and strongest in Churchill and Deception Bay (R^2 = .13 and .23, respectively; p < .05; Figure 6; Table S6), whereas a significant, albeit weaker, correlation was found in Iqaluit (R^2 = .09; p = .02; Table S6; Figure 6). For species collections, the correlation between beta diversity and geographic distance varied by port and collection method (zooplankton tow nets vs. benthos sampling methods). In Churchill, none of the correlations were significant (zooplankton R^2 = .014; p = .2, benthos R^2 = .004; p = .5; Table S6; Figure 6). For Deception Bay, a lower positive and significant correlation was found for the benthos (R^2 = .12, p = .02; Table S6; Figure 6) than for eDNA (R^2 = .23; p = .01; Table S6; Figure 6), while a stronger and significant positive correlation was found for zooplankton (R^2 = .26; p = .01; Table S6; Figure 6). For Iqaluit, a stronger and significant positive correlation was observed for the benthos (R^2 = .14, p = .01; Table S6; Figure 6) than for eDNA (R^2 = .09; p = .02; Table S6; Figure 6), while a negative and nonsignificant correlation was found for the zooplankton (R^2 = −.16; p > .05; Table S6; Figure 6).

3.5 | Origin of coastal eDNA

Taxa with the meroplanktonic life histories were the most commonly observed group based on eDNA sampling across ports (≥70% of observed taxa; Figure 7). Although the relative abundance of taxa by life history type varied among ports (PERMANOVA, p < .001), the proportions of taxa with meroplanktonic or holoplanktonic (taxa with only pelagic stage) life history types detected by eDNA were similar (Churchill: 69% meroplankton, 14% holoplankton; Deception Bay: 72% meroplankton, 17% holoplankton; Iqaluit: 80% meroplankton, 12% holoplankton; Figure 7). *Annelida* was the most dominant phylum detected with a meroplankton life history type, followed by *Mollusca* and *Echinodermata* (SIMPER analysis; 45.8% and 15.7% for both latter two species, respectively), whereas *Arthropoda* (copepods) was the dominant phylum in the holoplankton across the three ports (SIMPER analysis; 81.1%). Interestingly, similar dominant taxa were identified for the meroplankton component of communities detected via eDNA and species collection approaches, with the exception of *Echinodermata* for eDNA, which was replaced by *Arthropoda* (mostly amphipods) in species collection samples (SIMPER analysis; *Annelida* 45.6%, *Arthropoda* 24.0%, and *Mollusca* 16.5%). For holoplankton, *Arthropoda* (copepods) was the dominant phylum for both eDNA and zooplankton tows (SIMPER analysis; 81.1% and 96.1%, respectively).

4 | DISCUSSION

Arctic coastal regions are subject to harsh conditions, a wide range of temperatures and photoperiods, and support various forms of life over long periods of sea ice cover (PAME, 2016; Payne, Reusser, & Lee, 2012). Despite this, the Arctic Ocean is home to a great diversity of organisms, one which deserves increased attention, especially

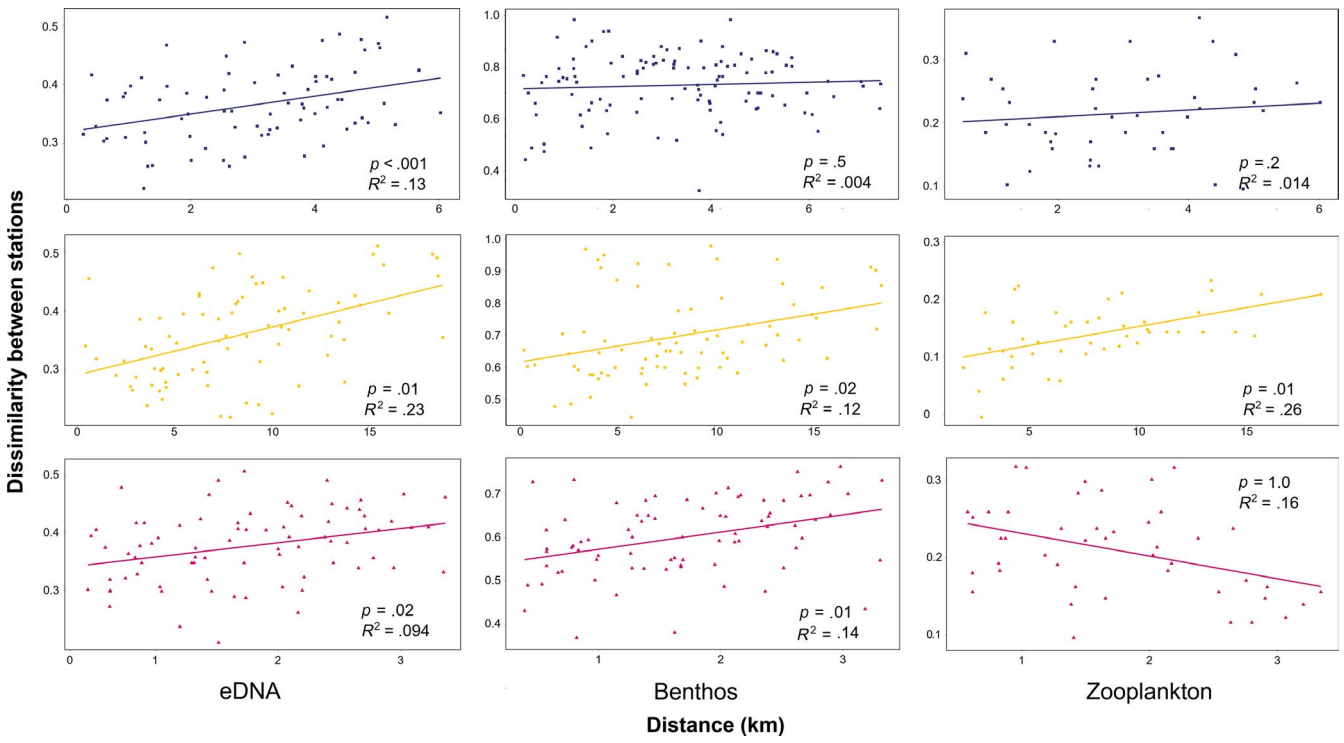


FIGURE 6 The Sorensen dissimilarity index between pairs of stations as a function of distance between the stations based on incidence data (presence/absence transformation on abundance) for different sampling methods (eDNA and species collections of benthos and zooplankton) in Churchill (blue), Deception Bay (yellow), and Iqaluit (magenta)

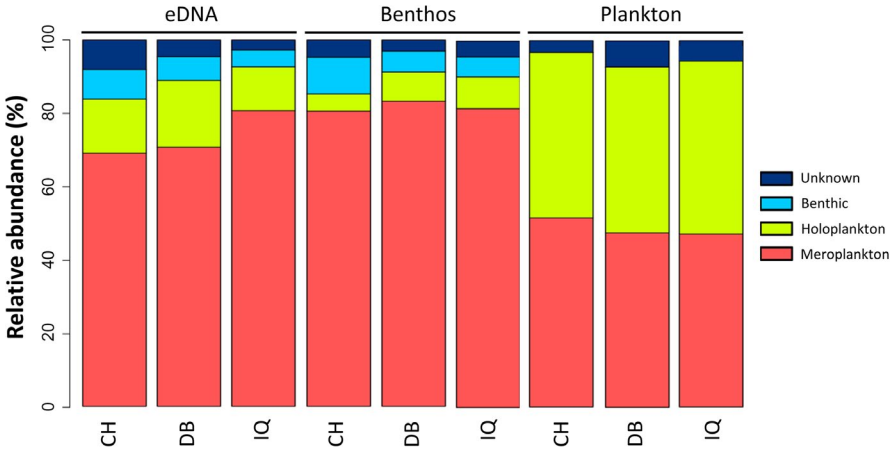


FIGURE 7 Relative abundance of organisms obtained with eDNA and species collection within Churchill, Deception Bay, and Iqaluit ports by life history type. Species collection for *benthos* includes benthic trawls, Van Veen grabs, and cores; *plankton* includes vertical and oblique pelagic plankton net tows. The sum of the detections for each genus (i.e., presence/absence) has been combined for all primer sets

lower trophic taxa, including invertebrates, which make up the base of ecosystem (Archambault et al., 2010; Piepenburg et al., 2011). The presence of marine invertebrates in the diets of Arctic fishes, birds, and mammals highlights their trophic importance (Bluhm & Gradinger, 2008; CAFF International Secretariat, 2010; Gajbhiye, 2002). Significant changes in their communities could thus affect ecosystem stability and impact the availability of food resources for coastal human communities (Guyot, Dickson, Paci, Furgal, & Chan, 2006; Ruiz, Carlton, Grosholz, & Hines, 1997). Marine biodiversity conservation is progressively becoming a crucial aim of environmental management (Spalding et al., 2007) but requires sufficient spatial data on biodiversity (Laurila-Pant et al., 2015). Despite substantial research efforts in recent years (Goldsmit et al., 2014; Piepenburg et al., 2011), there is limited knowledge about the diversity of many invertebrate groups (Archambault et al., 2010), including spatial distributions and how they are influenced by life stage transitions. Indeed, many species unknown to science await discovery (Jabr, Archambault, & Cameron, 2018; López, Olivier, Grant, & Archambault, 2016).

To our knowledge, this study is the first to compare eDNA, benthos, and zooplankton community patterns in the Arctic. Our use of eDNA sampling in parallel with species collection at Arctic ports provides insight into the ecological properties of eDNA in relation to the distribution and life stages of coastal marine invertebrates. While differing from observations made using species collection approaches, eDNA metabarcoding of Arctic coastal zone taxa provided relevant, complementary biodiversity information at various spatial scales using alpha, beta, and gamma indices.

4.1 | Overall biodiversity and community structure

Despite limited sample volumes (only 30 L water in total) and sequencing depth, eDNA metabarcoding identified 202 marine genera, covering 15 phyla and complementing biodiversity information obtained from species collection using traditional benthic trawls, cores, grabs, and net tows, representing a combined total of 634 genera, covering 23 phyla for eDNA and species collection. Following the qualitative results obtained by Thomsen et al. (2016) when comparing fish biodiversity detected by eDNA and species caught by trawl offshore Greenland, a greater similarity between sampling methods was expected. Instead, we observed important differences between phylum whereby *Bryozoa*, *Arthropoda*, and *Mollusca* were more commonly encountered with species collections of coastal marine communities while *Echinodermata*, *Porifera*, *Nemertea*, and *Cnidaria* were more frequently detected in eDNA samples. Several physical and biological factors might explain the differences in detectability of taxa between approaches. For example, echinoderms and sponges (*Porifera*) are often attached to large boulders in the seabed (Bell & Barnes, 2003; Chapman, 2003) and are difficult to collect using trawls or grabs, which may negatively bias their detectability in species-based collections. Identification issues, directly or in combination with biases in detectability, may also explain differences in community assemblages identified through eDNA and species collections. For instance, ribbon worms often lack easily diagnosable external body features making identification challenging and are frequently found under rocks, making them difficult to access (Thiel & Norenburg, 2009). eDNA metabarcoding may thus be particularly useful in such cases where taxa are more difficult to sample or identify morphologically. It is also important to note the considerable phylum-specific variation in previous sequencing efforts which impacts the chance of eDNA from a given group of being matched to sequences of morphologically identified organisms. For example, 54.5%–56.3% of the *Arthropods*, *Cnidarians*, and *Mollusks* identified by our traditional collection sampling methods were present in the sequence databases, while only 28.6% of the bryozoans had been previously sequenced for the barcoding regions used in this study (Table S7). This clearly limits the ability of eDNA metabarcoding to fully document community composition in the Arctic and highlights the importance of improving sequencing effort for particular taxa to fill the taxonomic gaps in available databases.

Another salient observation of this study is that detected community structure differed substantially between sampling methods with benthic communities being more variable within and between harbors and zooplankton communities being more similar within and between harbors. The broader range of biodiversity dissimilarities observed among benthic communities may be explained by highly

variable seabed characteristics, which play an important role in distribution of megafauna as they impact several factors, including larval settlement, anchorages, and shelter (Kedra, Renaud, Andrade, Goszczko, & Ambrose, 2013; Preez, Curtis, & Clarke, 2016). In contrast, zooplankton experience less variation in their habitat, due to the greater homogeneity of the water column relative to benthic substrates (Angel, 1993; Gray, 1997). Variation in eDNA community structure was intermediate between the variation observed using the two different species collection approaches. Thus, eDNA community structure represented greater community dissimilarity than what was observed for plankton communities but less dissimilarity than what was observed for benthic communities (trawl, grabs, and cores). This pattern could be due to the origin of eDNA, transport, and degradation processes. The high prevalence of meroplanktonic organisms (reflective of benthic communities) detected within eDNA communities may explain why they display greater dissimilarity than do plankton communities as depicted by species collections. On the other hand, eDNA communities likely display less dissimilarity than do benthic communities as depicted by species collections due to the homogenization and degradation of eDNA particle in the water column, whereas living specimens remain in/on seafloor and are less affected by water movement. In the future, it would be relevant to characterize habitats from which the samples originate to see if the eDNA approach could have detected differences in microhabitats, for instance, as reported by Port et al. (2016). Similarly, as the biological substrate sampled for eDNA is a critical factor influencing the biotic composition (Hermans, Buckley, & Lear, 2018; Koziol et al., 2019), the use of eDNA sediment substrates in addition to the eDNA water samples might have revealed dissimilarity patterns closer to the benthic communities. Our observations of distinct patterns of community structure depicted using either COI and 18S primer sets are consistent with several studies that have shown an effect of markers on the detection rate of marine invertebrates (Djurhuus et al., 2018; Drummond et al., 2015; Elbrecht et al., 2017; Kelly et al., 2017; Shaw et al., 2016). This highlights the importance of using a combination of different primer sets covering different genomic regions until a more universal primer set is available. Here, our results suggested a greater affinity of COI primers for *Annelids*, *Arthropods*, and *Echinoderms* relative to 18S primers, as previously reported by Drummond et al. (2015). These affinities could potentially explain why the observed Iqaluit community composition based on COI and 18S clearly differed from Churchill and Deception Bay communities as more *Annelids* and *Echinoderms* and less *Arthropods* taxa were detected in Iqaluit relative to the other two locations.

Despite the large number of taxa observed in this study, many marine invertebrates were likely missed, as suggested by the rarefaction curves. This is especially true for benthic communities, for which the rarefaction curves showed little indication that species increases were slowing. Coastal areas present complex mosaics of benthic habitat which, in addition to creating diverse epi- and infaunal communities, increases the possibility of missing taxa when sampling (Gray, 1997). For eDNA sampling, the number of genera detected may be influenced by a number of factors, including sample size and their vertical and horizontal distributions (Lacoursière-Roussel et al., 2018), filter types, volume of filtered water, extraction method (Deiner et al., 2018), sequencing depth, and bioinformatics pipeline. Thus, a larger volume of filtered seawater for each sample (Shaw et al., 2016) and a greater sequencing depth would likely have improved the detection rate (Mächler, Deiner, Spahn, & Altermatt, 2016) and increased the observed generic richness. Similarly, a greater detection rate could have been achieved by sampling a greater number of stations within each port. Although eDNA rarefaction curves were very similar between Churchill and Deception Bay harbors, Iqaluit grew less rapidly at first and appeared closer to reaching a plateau than did Churchill and Deception Bay due to the lower alpha and gamma biodiversity measured with this harbor. Further, alpha biodiversity and gamma biodiversity were greater within Iqaluit for species collections. This suggests that the opposing trends observed for the two approaches might reflect decreased previous monitoring effort in more northern regions which would logically result in more incomplete sequence reference databases rather than a true lower biodiversity. Sequence reference databases are estimated to contain only 13% of marine species inhabiting the Arctic Ocean (Hardy et al., 2011), and a latitudinal gradient of sequencing effort might exist within the Arctic itself. Indeed, we observed an increasing fraction of unknown OTUs from Churchill north to Iqaluit.

4.2 | Transport and homogenization of eDNA

Knowledge on the spatial arrangement of biodiversity is crucial for protecting regional diversity and supporting conservation planning (Socolar et al., 2015). The complex mosaic of benthic habitats in Arctic coastal areas makes it difficult to obtain a comprehensive sampling of this component of biodiversity. Our results found much lower beta diversity for eDNA communities compared to species collection communities which, suggesting that species eDNA is more homogeneous in space than the associated species themselves in coastal zones, as has been observed in several studies of freshwater systems (Dejean et al., 2011; Ficetola et al., 2008; Li et al., 2018; Thomsen et al., 2012).

Although Arctic coastal eDNA showed a more homogeneous community structure than do the composite species, this pattern was affected by spatial scale. Indeed, our results revealed a significant relationship between the dissimilarities within eDNA communities as a function of geographic distance, spanning distances from 4 to nearly 20 km. This is consistent with many spatial ecology processes whereby communities close to one another are more similar than are those that are further apart (Nekola & White, 1999), and in line with the observations of O'Donnell et al. (2017) of greater eDNA dispersion in nearshore marine habitats. Several studies have also revealed patterns of extensive eDNA dispersion over considerable distances within river systems (Deiner & Altermatt, 2014; Deiner, Fronhofer, Mächler, Walser, & Altermatt, 2016), which could influence community structure in estuarine settings such as the port of Churchill. In our study, the very cold Arctic waters may further contribute to reducing DNA degradation, thus

providing more time for dispersion over larger distances compared to what has been previously reported at more temperate latitudes (Jeunen et al., 2019). This raises the hypothesis that spatial eDNA homogenization should be more important in the Arctic Ocean than more southern regions. In contrast, given that sunlight is known to break down DNA in marine systems (El-Sayed, Van Dijken, & Gonzalez-Rodas, 1996), the prolonged daylight in the study sites at the time of sampling (up to 24 hr) may encourage DNA degradation (Mächler, Osathanunkul, & Altermatt, 2018). However, a study by Andruszkiewicz, Sassoubre, and Boehm (2017) concluded that sunlight may not be the primary factor causing degradation of the fish DNA in their experiment and that degradation of the latter would depend more on the time elapsed since its shedding in the water. As many chemical and biological processes influence eDNA production, transport, and degradation, it will be of interest in future studies to evaluate how latitude may influence patterns of eDNA biodiversity indices.

The weak correlation between dissimilarity and geographic distance in Iqaluit is in sharp contrast to the other two ports in this study. This may be explained by the greater tidal range in the region (7.5–11.7 m, as compared to 3.3–5.1 and 3.6–5.7 m for Churchill and Deception Bay, respectively) and associated currents occurring in this location (Fisheries & Oceans Canada, 2018). Interestingly, Churchill and Deception Bay ports showed significant distance differences between their stations (Churchill: 0.2–7 km; Deception Bay: 0.3–19 km), suggesting that the correlation between dissimilarity and distance might be consistent at various spatial scales for marine invertebrates in Arctic coastal environments with similar tidal conditions. In contrast to eDNA results, where dissimilarity increased as a function of geographic distance between stations, increased dissimilarity of communities with distance was not systematically observed in species collections, which again may reflect the fact that marine invertebrate communities are often characterized by a pronounced patchiness (Ministry of Environment, 2006). Thus, the homogeneity of eDNA distribution due to dispersion could potentially improve estimations of biodiversity at local spatial scales. On the other hand, the dispersion and persistence of eDNA in coastal environments also increase the risk of detecting organisms that are not actually present locally (Deiner & Altermatt, 2014; Jane et al., 2015). Further studies comparing the spatial distribution of eDNA communities and corresponding species collection communities (either benthos or plankton) in dynamic systems such as complex coastal areas are needed to improve our knowledge about how the multiple physical and biological factors influence eDNA distance decay. Such information will help to better inform eDNA sampling design for monitoring and management issues.

4.3 | Origins of eDNA

Benthic species with meroplanktonic life history type accounted for a greater proportion of the eDNA than did species with strict benthic or pelagic life history. This result suggests that coastal water eDNA is a mixture of organic material released to the environment (e.g., feces, skin, mucus) and plankton degradation and thus underlines the influence of variation in the life cycles on species detection probability. For instance, the fact that the discriminating taxa collected using eDNA and species collection approaches differed for holoplankton and meroplankton communities suggests that the different reproductive periods of the organisms, as well as the associated planktonic larval stages, may influence the detection of certain taxa. As a case in point, the daisy brittle star (*Ophiopholis aculeata*), the brittle star *Ophiura robusta*, and the green sea urchin (*Strongylocentrotus droebachiensis*) were discriminant echinoderm species detected by eDNA and not by benthic species collection (data not shown). Interestingly, these three species are known to synchronize their spawning periods with sharp increases in sea temperature (Himmelman, Dumont, Gaymer, Vallières, & Drolet, 2008), which typically occur during July within the sampled ports (Galbraith & Larouche, 2011; Prinsenberg, 1984), suggesting that the high number of sequences observed for those species could reflect the occurrence of these species in their pelagic phase.

The importance of planktonic stages to increasing eDNA detection is also supported by the absence of DNA from *Amphipods*, which were discriminant taxa in species collections for meroplankton. In general, studies on amphipod reproductive biology revealed that breeding occurs during the spring in most species (Węsławski & Legeżyńska, 2002). However, amphipods represent a complex case as some species are benthic while other species are planktonic and the two life history types coexist in the same environment. Sampling outside of breeding periods and the lack of a planktonic stage could explain the lower detectability of these organisms with eDNA. It is difficult to draw general patterns based on the life histories of organisms since species or genera differ substantially and there is a general lack of knowledge for life histories, including their reproduction periods, of many marine invertebrates inhabiting the Arctic. O'Donnell et al. (2017) also concluded that planktonic larval stages or released pelagic eggs may play an important role in the eDNA detection of some organisms. However, given that seasonal factors greatly influence the proportion of meroplanktonic and holoplanktonic organisms (Highfield et al., 2010; Lindeque, Parry, Harmer, Somerfield, & Atkinson, 2013) and eDNA ecology (e.g., water temperature, UV exposition), further studies on the detection of various marine invertebrate taxa at different times of the year would aid to determine how life histories of different organisms impact eDNA detection.

4.4 | Role of eDNA in Arctic conservation

Given the multiple environmental and anthropogenic factors that are currently threatening Arctic coastal biodiversity and the international objectives that many nations have agreed to, such as the protection of 10% of coastal and marine areas by 2020 (Secretariat of the Convention on Biological Diversity [SCBD], 2014), the development of rapid and efficient tools for monitoring biodiversity

changes is essential. eDNA metabarcoding provides valuable information toward a broader view of the taxonomic diversity that may help in developing more rigorous conservation plans, particularly in the Arctic. In addition, this approach provides numerous advantages due to its time-efficient and nonintrusive nature (Deiner et al., 2017). The simplicity of the sampling protocol for coastal water makes the method easy to learn, which constitutes a major asset for remote regions such as the Arctic, where it can be easily incorporated into existing sampling or community-based monitoring programs (Lacoursière-Roussel et al., 2018). By combining the study of invertebrate communities at different spatial scales detected by eDNA and species collection, this study highlights important features related to the ecology of eDNA biodiversity indices such as the origin of eDNA (i.e., planktonic phases of benthic taxa) and the effect of spatial homogenization. Together, our results suggest that eDNA diversity reflects complex interactions between the life cycles of organisms and their spatial distribution. As public sequence databases become more complete over time, species detection using eDNA metabarcoding will improve and is likely to increase understanding of a wide range of ecological processes (daily plankton migration, seasonal fish migration, food web interactions, etc.) where many elements remain undiscovered. Our results highlight that eDNA should be used as a complementary approach for improving characterization of coastal biodiversity from species collections as each method yielded distinct information on taxonomic composition of the invertebrates inhabiting coastal areas.

ACKNOWLEDGMENTS

This project was funded by ArcticNet, POLAR Knowledge, Fisheries and Oceans Canada Aquatic Invasive Species Monitoring Program, Nunavut Wildlife Management Board, and Nunavik Marine Region Wildlife Board. We thank Brian Boyle from IBIS for his expertise at the sequencing platform; David Lodge, Kristy Deiner, and Erin K Grey for their help and advice in terms of eDNA metabarcoding methodologies and analyses; and Jésica Goldsmit for her essential knowledge on shipping activities and NIS in the Arctic and sharing benthic core data from the Port of Churchill. We thank Melania Cristescu and Guang Zhang for their help in primer selection and Frederic Chain and Yiyuan Li for the pipeline development, Cecilia Hernandez for laboratory assistance, and Jérôme Laroche from IBIS for the development of bioinformatics pipelines. We would like to thank Laure de Montety, taxonomist, and members of the Archambault laboratory for the identification of the benthic organisms as well as for their knowledge about them. We gratefully acknowledge the Churchill Northern Studies Centre and Glencore Raglan Mine for providing access to their building facilities. We also thank the following individuals for field assistance and participating in training: Frédéric Hartog, Valérie Cypihot, Cyndy Grant, LeeAnn Fishback, Daniel Gibson, Dick Hunter, Austin MacLeod, Thomas Whittle, Rory McDonald, Frederic Lemire, Adamie Keatanik, Willie Keatanik, Willie Alaku, and Markusie Jaaka.

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

AL-R, KLH, PA, and LB conceived the study. NL, AL-R, KLH, CWM, NS, and AD contributed to the data acquisition in the field. KLH, PA, CWM, and NS are specialized in the Arctic coastal surveillance and contributed to benthic component of the study dataset. GW and AD are specialized in Arctic zooplankton monitoring and contributed to the zooplankton component of the study dataset. EN developed the bioinformatics pipeline. NL, AL-R, MS, PA, and LB interpreted the data. NL wrote the manuscript, and all authors reviewed it.

DATA AVAILABILITY STATEMENT

Raw sequence reads were deposited in NCBI's Sequence Read Archive (SRA, http://www.ncbi.nlm.nih.gov/sra) under Bioprojects PRJNA388333 and PRJNA521343. The data that support the findings of this study are openly available in [repository name e.g., "figshare"] at http://doi.org/10.1002/edn3.35, reference number [16575833].

ORCID

Noémie Leduc https://orcid.org/0000-0002-0778-144X
Louis Bernatchez https://orcid.org/0000-0002-8085-9709

REFERENCES

ACIA (2004). *Impacts of a warming arctic; Arctic climate impact assessment*. Cambridge, UK: Cambridge University Press.

Andruszkiewicz, E. A., Sassoubre, L. M., & Boehm, A. B. (2017). Persistence of marine fish environmental DNA and the influence of sunlight. *PLoS ONE*, *12*(9), e0185043. https://doi.org/10.1371/journal.pone.0185043

Angel, M. V. (1993). Biodiversity of the Pelagic Ocean. *Conservation Biology*, *7*(4), 760–772. https://doi.org/10.1046/j.1523-1739.1993.740760.x

Archambault, P., Snelgrove, P. V. R., Fisher, J. A. D., Gagnon, J., David, J., Harvey, M., Pepin, P. (2010). From sea to sea: Canada's three oceans of biodiversity. *PLoS ONE*, *5*(8), e12182. https://doi.org/10.1371/journal.pone.0012182

Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, *17*(1), 1–17. https://doi.org/10.1007/s10592-015-0775-4

Bell, J. J., & Barnes, D. K. A. (2003). Effect of disturbance on assemblages: An example using *Porifera*. *Biological Bulletin*, *205*(2), 144–159. https://doi.org/10.2307/1543235

Bivand, R. S., Pebesma, E. J., & Gomez-Rubio, V. (2008). *Applied spatial data analysis with R*. Berlin, Germany: Springer. https://doi.org/10.1007/978-0-387-78171-6

Bluhm, B. A., & Gradinger, R. (2008). Regional variability in food availability for arctic marine mammals. *Ecological Applications*, *18*(Suppl 2), S77–S96. https://doi.org/10.1890/06-0562.1

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

CAFF International Secretariat (2010). *Arctic biodiversity trends 2010 – Selected indicators of change*. Akureyri, Iceland: CAFF International Secretariat.

CAFF International Secretariat (2013). *Arctic Biodiversity Assessment – Status and trends in Arctic biodiversity. Conservation of Arctic Flora and Fauna*. Akureyri, Iceland: CAFF International Secreteriat.

Carlton, J. T. (1996). Biological invasions and cryptogenic species. *Ecology*, *77*(6), 1653–1655. https://doi.org/10.2307/2265767

Casas-Monroy, O., Linley, R.D., Adams, J.K., Chan, F.T., Drake, D.A.R., & Bailey, S.A. (2014). National Risk Assessment for Introduction of Aquatic Nonindigenous Species to Canada by Ballast Water. DFO Can. Sci. Advis. Sec. Res. Doc. 2013/128. vi + 73 p.

Chan, F. T., Bailey, S. A., Wiley, C. J., & MacIsaac, H. J. (2013). Relative risk assessment for ballast-mediated invasions at Canadian Arctic ports. *Biological Invasions*, *15*(2), 295–308. https://doi.org/10.1007/s10530-012-0284-z

Chan, F. T., Bronnenhuber, J.E., Bradie, J.N., Howland, K., Simard, N., & Bailey, S.A. (2012). Risk assessment for ship-mediated introductions of aquatic nonindigenous species to the Canadian Arctic. DFO Can. Sci. Advis. Sec. Res. Doc. 2011/105. vi + 93 p.

Chapman, M. G. (2003). The use of sandstone blocks to test hypotheses about colonization of intertidal boulders. *Journal of the Marine Biological Association of the United Kingdom*, *83*(2), 415–423. https://doi.org/10.1017/S0025315403007276h

Clarke, K. R., & Gorley, R. N. (2006). *PRIMER V6: User Manual/Tutorial* (192 p). Plymouth, UK: PRIMER-E.

Clarke, K. R., & Warwick, R. M. (2001). *Changes in communities: An approach to statistical analysis and interpretation* (2nd ed.). Plymouth, UK: PRIMER-E.

Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, W. K., … Bik, H. M. (2016). The ecologist's field guide to sequence-based identification of biodiversity. *Methods in Ecology and Evolution*, *7*(9), 1008–1018. https://doi.org/10.1111/2041-210X.12574

Cusson, M., Archambault, P., & Aitken, A. (2007). Biodiversity of benthic assemblages on the Arctic continental shelf : Historical data from Canada. *Marine Ecology Progress Series*, *331*, 291–304.

Darnis, G., Robert, D., Pomerleau, C., Link, H., Archambault, P., Nelson, R. J., … 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. https://doi.org/10.1007/s10584-012-0483-8

Deiner, K., & Altermatt, F. (2014). Transport distance of invertebrate environmental DNA in a natural river. *PLoS ONE*, *9*(2), e88786. https://doi.org/10.1371/journal.pone.0088786

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., … Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, *26*(21), 5872–5895. https://doi.org/10.1111/mec.14350

Deiner, K., Fronhofer, E. A., Mächler, E., Walser, J. C., & Altermatt, F. (2016). Environmental DNA reveals that rivers are conveyer belts of biodiversity information. *Nature Communications*, *7*, 1–9. http://doi.org/10.1038/ncomms12544

Deiner, K., Lopez, J., Bourne, S., Holman, L. E., Seymour, M., Grey, E. K., … Lodge, D. M. (2018). Optimising the detection of marine taxonomic richness using environmental DNA metabarcoding : the effects of filter material , pore size and extraction method. *Metabarcoding and Metagenomics*, *2*, 1–15. http://doi.org/10.3897/mbmg.2.28963

Dejean, T., Valentini, A., Duparc, A., Pellier-Cuit, S., Pompanon, F., Taberlet, P., & Miaud, C. (2011). Persistence of environmental DNA in freshwater ecosystems. *PLoS ONE*, *6*(8), 8–11. https://doi.org/10.1371/journal.pone.0023398

Dispas, A. (2019). *Étude de référence sur la biodiversité du mésozooplancton dans quatre ports de l'Arctique canadien en vue d'une augmentation de l'activité maritime, de l'exploitation des ressources et du réchauffement climatique* (146 p). Québec, QC: Université du Québec à Rimouski.

Djurhuus, A., Pitz, K., Sawaya, N. A., Rojas-Márquez, J., Michaud, B., Montes, E., Breitbart, M. (2018). Evaluation of marine zooplankton community structure through environmental DNA metabarcoding. *Limnology and Oceanography: Methods*, *16*, 209–221. https://doi.org/10.1002/lom3.10237

Dray, S., & Dufour, A.-B. (2007). The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, *22*(4), 1–20.

Drummond, A.. J., Newcomb, R. D., Buckley, T. R., Xie, D., Dopheide, A., Potter, B. C. M., … Nelson, N. (2015). Evaluating a multigene environmental DNA approach for biodiversity assessment. *GigaScience*, *4*(46), 1–19. http://doi.org/10.1186/s13742-015-0086-1

Elbrecht, V., & Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass-sequence relationships with an innovative metabarcoding protocol. *PLoS ONE*, *10*(7), 1–16. https://doi.org/10.1371/journal.pone.0130324

Elbrecht, V., Vamos, E. E., Meissner, K., Aroviita, J., & Leese, F. (2017). Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution*, *8*(10), 1265–1275. https://doi.org/10.1111/2041-210X.12789

El-Sayed, S. Z., Van Dijken, G. L., & Gonzalez-Rodas, G. (1996). Effects of ultraviolet radiation on marine ecosystems. *International Journal of Environmental Studies*, *51*(3), 199–216. https://doi.org/10.1080/00207239608711081

Ficetola, G. F., Miaud, C., Pompanon, F., & Taberlet, P. (2008). Species detection using environmental DNA from water samples. *Biology Letters*, *4*(4), 423–425. https://doi.org/10.1098/rsbl.2008.0118

Fisheries and Oceans Canada (2018). *Canadian tide and current tables*. Retrieved from http://www.charts.gc.ca/publications/tables-eng.asp

Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, *3*(5), 294–299. https://doi.org/10.1371/journal.pone.0013102

Gajbhiye, S. N. (2002). Zooplankton study methods, Importance and significant observations. *The National Seminar on Creeks, Estuaries and Mangroves - Pollution and Conservation*. (pp. 21–27).

Galbraith, P. S., & Larouche, P. (2011). Sea-surface temperature in Hudson Bay and Hudson Strait in relation to air temperature and ice cover breakup, 1985–2009. *Journal of Marine Systems*, *87*(1), 66–78. https://doi.org/10.1016/j.jmarsys.2011.03.002

Geller, J., Meyer, C., Parker, M., & Hawk, H. (2013). Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Molecular Ecology Resources*, *13*(5), 851–861. https://doi.org/10.1111/1755-0998.12138

Goldsmit, J. (2016). *Benthic non-indigenous species in Ports of the Canadian Arctic: Identification, biodiversity and relationships with global warming and shipping activity* (251 p.). Doctoral Thesis. Québec, QC: Université du Québec à Rimouski.

Goldsmit, J., Archambault, P., Chust, G., Villarino, E., Liu, G., Lukovich, J. V., … Howland, K. L. (2018). Projecting present and future habitat suitability of ship- mediated aquatic invasive species in the Canadian Arctic. *Biological Invasions*, *20*(2), 501–517. https://doi.org/10.1007/s10530-017-1553-7

Goldsmit, J., Archambault, P., & Howland, K. L. (2014). Establishing a baseline for early detection of non-indigenous species in ports of the Canadian Arctic. *Aquatic Invasions*, *9*(3), 327–342. https://doi.org/10.3391/ai.2014.9.3.08

Goldsmit, J., McKindsey, C., Archambault, P., & Howland, K. L. (2019). Ecological risk assessment of predicted marine invasions in the Canadian Arctic. *PLoS ONE*, *14*(2), e0211815. https://doi.org/10.1371/journal.pone.0211815

Gradinger, R., Bluhm, B. A., Hopcroft, R. R., Gebruk, A. V., Kosobokova, K., Sirenko, B., & Weslawski, J. M. (2010). Marine life in the Arctic. InA. D. McIntyre (Ed.), *Life in the World's oceans: Diversity, distribution, and abundance* (pp. 183–202). Hoboken, NJ: Blackwell Publishing, Ltd. https://doi.org/10.1002/9781444325508.ch10

Gray, J. S. (1997). Marine biodiversity: Patterns, threats and conservation needs. *Biodiversity and Conservation*, *6*(1), 153–175. https://doi.org/10.1023/A:1018335901847

Grey, E. K., Bernatchez, L., Cassey, P., Deiner, K., Deveney, M., Howland, K. L., … Lodge, D. M. (2018). Effects of sampling effort on biodiversity patterns estimated from environmental DNA metabarcoding surveys. *Scientific Reports*, *8*(1), 2–11. https://doi.org/10.1038/s41598-018-27048-2

Guyot, M., Dickson, C., Paci, C., Furgal, C., & Chan, H. M. (2006). Local observations of climate change and impacts on traditional food security in two northern Aboriginal communities. *International Journal of Circumpolar Health*, *65*(5), 403–415. https://doi.org/10.3402/ijch.v65i5.18135

Hadziavdic, K., Lekang, K., Lanzen, A., Jonassen, I., Thompson, E. M., & Troedsson, C. (2014). Characterization of the 18S rRNA gene for designing universal eukaryote specific primers. *PLoS ONE*, *9*(2), e87624. https://doi.org/10.1371/journal.pone.0087624

Hansen, B. K., Bekkevold, D., Clausen, L. W., & Nielsen, E. E. (2018). The sceptical optimist: Challenges and perspectives for the application of environmental DNA in marine fisheries. *Fish and Fisheries*, *19*, 751–768. https://doi.org/10.1111/faf.12286

Hardy, S. M., Carr, C. M., Hardman, M., Steinke, D., Corstorphine, E., & Mah, C. (2011). Biodiversity and phylogeography of Arctic marine fauna: Insights from molecular tools. *Marine Biodiversity*, *41*(1), 195–210. https://doi.org/10.1007/s12526-010-0056-x

Hermans, S. M., Buckley, H. L., & Lear, G. (2018). Optimal extraction methods for the simultaneous analysis of DNA from diverse organisms and sample types. *Molecular Ecology Resources*, *18*, 557–569. https://doi.org/10.1111/1755-0998.12762

Highfield, J. M., Eloire, D., Conway, D. V. P., Lindeque, P. K., Attrill, M. J., & Somerfield, P. J. (2010). Seasonal dynamics of meroplankton assemblages at station L4. *Journal of Plankton Research*, *32*(5), 681–691. https://doi.org/10.1093/plankt/fbp139

Himmelman, J. H., Dumont, C. P., Gaymer, C. F., Vallières, C., & Drolet, D. (2008). Spawning synchrony and aggregative behaviour of cold-water echinoderms during multi-species mass spawnings. *Marine Ecology Progress Series*, *361*, 161–168. https://doi.org/10.3354/meps07415

Intergovernmental Panel on Climate Change (IPCC) (2018). Summary for policymakers. In *Global warming of 1.5°C. An IPCC special report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty* (32 pp). Geneva, Switzerland: World Meteorological Organization.

IOC/UNESCO (2010). *Why monitor the Arctic Ocean? Services to society from sustained ocean observing system*. Retrieved from: http://unesdoc.unesco.org/images/0018/001898/189843e.pdf

Jabr, N., Archambault, P., & Cameron, C. (2018). Biogeography and adaptations of torquaratorid acorn worms (Hemichordata: Enteropneusta) including two new species from the Canadian Arctic. *Canadian Journal of Zoology*, *96*(11), 1221–1229. https://doi.org/10.1139/cjz-2017-0214

Jane, S. F., Wilcox, T. M., McKelvey, K. S., Young, M. K., Schwartz, M. K., Lowe, W. H., … Whiteley, A. R. (2015). Distance, flow and PCR inhibition: EDNA dynamics in two headwater streams. *Molecular Ecology Resources*, *15*(1), 216–227. https://doi.org/10.1111/1755-0998.12285

Jarosław, K., Mioduchowska, M., & Petković, M. (2016). Trying to solve current issues with invertebrate taxonomy – The conceptual web-based application. *World Scientific News*, *57*, 664–673.

Jeunen, G.-J., Knapp, M., Spencer, H. G., Lamare, M. D., Taylor, H. R., Stat, M., … Gemmell, N. J. (2019). Environmental DNA (eDNA) metabarcoding reveals strong discrimination among diverse marine habitats connected by water movement. *Molecular Ecology Resources*, *19*, 426–438. https://doi.org/10.1111/1755-0998.12982

Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., … Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, *16*(10), 1245–1257. https://doi.org/10.1111/ele.12162

Jones, J. B. (1992). Environmental impact of trawling on the seabed: A review. *New Zealand Journal of Marine and Freshwater Research*, *26*(1), 59–67. https://doi.org/10.1080/00288330.1992.9516500

Jørgensen, L. L., Archambault, P., Piepenburg, D., & Rice, J. (2016). *Arctic marine biodiversity*. Retrieved from https://www.researchgate.net/publication/292115665_Arctic_marine_biodiversity

Kedra, M., Renaud, P. E., Andrade, H., Goszczko, I., & Ambrose, W. G. (2013). Benthic community structure, diversity, and productivity in the shallow Barents Sea bank (Svalbard Bank). *Marine Biology*, *160*(4), 805–819. https://doi.org/10.1007/s00227-012-2135-y

Kelly, R. P., Closek, C. J., O'Donnell, J. L., Kralj, J. E., Shelton, A. O., & Samhouri, J. F. (2017). Genetic and manual survey methods yield different and complementary views of an ecosystem. *Frontiers in Marine Science*, *3*, 1–11. https://doi.org/10.3389/fmars.2016.00283

Kendall, M. A. (1996). Are Arctic soft sediment macrobenthic communities impoverished? *Polar Biology*, *16*(6), 393–399. https://doi.org/10.1007/s003000050070

Knowlton, N. (1993). Sibling species in the sea. *Annual Review of Ecology and Systematics*, *24*(1), 189–216. https://doi.org/10.1146/annurev.ecolsys.24.1.189

Koziol, A., Stat, M., Simpson, T., Jarman, S., DiBattista, J. D., Harvey, E. S., … Bunce, M. (2019). Environmental DNA metabarcoding studies are critically affected by substrate selection. *Molecular Ecology Resources*, *19*, 366–376. https://doi.org/10.1111/1755-0998.12971

Kwong, S., Srivathsan, A., & Meier, R. (2012). An update on DNA barcoding : Low species coverage and numerous. *Cladistics*, *28*, 639–644.

Lacoursière-Roussel, A., & Deiner, K. (in press). 'Environmental DNA' is not the tool by itself. *Journal of Fish Biology*.

Lacoursière-Roussel, A., Howland, K., Normandeau, E., Grey, E. K., Archambault, P., Deiner, K., … Bernatchez, L. (2018). eDNA metabarcoding as a new surveillance approach for coastal Arctic biodiversity. *Ecology and Evolution*, *8*, 7763–7777. https://doi.org/10.1002/ece3.4213

Laurila-Pant, M., Lehikoinen, A., Uusitalo, L., & Venesjärvi, R. (2015). How to value biodiversity in environmental management? *Ecological Indicators*, *55*, 1–11. https://doi.org/10.1016/j.ecolind.2015.02.034

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., … Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, *10*(1), 34. https://doi.org/10.1186/1742-9994-10-34

Li, Y., Evans, N. T., Renshaw, M. A., Jerde, C. L., Olds, B. P., Shogren, A. J., … Pfrender, M. E. (2018). Estimating fish alpha- and beta-diversity along a small stream with environmental DNA metabarcoding. *Metabarcoding and Metagenomics*, *2*, 1–11. http://doi.org/10.3897/mbmg.2.24262

Lindeque, P. K., Parry, H. E., Harmer, R. A., Somerfield, P. J., & Atkinson, A. (2013). Next generation sequencing reveals the hidden diversity of zooplankton assemblages. *PLoS ONE*, *8*(11), 1–14. https://doi.org/10.1371/journal.pone.0081327

Link, H., Chaillou, G., Forest, A., Piepenburg, D., & Archambault, P. (2013). Multivariate benthic ecosystem functioning in the Arctic-benthic fluxes explained by environmental parameters in the south-eastern Beaufort Sea. *Biogeosciences*, *10*(9), 5911–5929. https://doi.org/10.5194/bg-10-5911-2013

López, E., Olivier, F., Grant, C., & Archambault, P. (2016). Four new records and a new species of sedentary polychaetes from the Canadian High Arctic. *Journal of the Marine Biological Association of the United Kingdom*, *97*(8), 1685–1695. https://doi.org/10.1017/S0025315416000953

MacArthur, R. H. (1965). Patterns of species diversity. *Biological Reviews*, *40*(4), 510–533. https://doi.org/10.1111/j.1469-185X.1965.tb00815.x

Mächler, E., Deiner, K., Spahn, F., & Altermatt, F. (2016). Fishing in the water: Effect of sampled water volume on environmental DNA-based detection of macroinvertebrates. *Environmental Science & Technology*, *50*(1), 305–312. https://doi.org/10.1021/acs.est.5b04188

Mächler, E., Osathanunkul, M., & Altermatt, F. (2018). Shedding light on eDNA: Neither natural levels of UV radiation nor the presence of a filter feeder affect eDNA-based detection of aquatic organisms. *PLoS ONE*, *13*(4), e0195529. https://doi.org/10.1371/journal.pone.0195529

Magoč, T., & Salzberg, S. L. (2011). FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, *27*(21), 2957–2963. https://doi.org/10.1093/bioinformatics/btr507

Marcus, N. H., & Boero, F. (1998). Minireview: The importance of benthic-pelagic coupling and the forgotten role of life cycles in coastal aquatic systems. *Limnology and Oceanography*, *43*(5), 763–768. https://doi.org/10.4319/lo.1998.43.5.0763

McGill, B. J., Dornelas, M., Gotelli, N. J., & Magurran, A. E. (2015). Fifteen forms of biodiversity trend in the Anthropocene. *Trends in Ecology and Evolution*, *30*, 104–113. https://doi.org/10.1016/j.tree.2014.11.006

Millennium Ecosystem Assessment (2005). *Ecosystems and human well-being: Synthesis*. Island Press. Retrieved from www.islandpress.org

Ministry of Environment (2006). *Guidelines for sampling benthic invertebrates in British Columbia Streams*. Vancouver, BC: Ministry of Environment.

Mori, A. S., Isbell, F., & Seidl, R. (2018). β-Diversity, community assembly, and ecosystem functioning. *Trends in Ecology and Evolution*, *33*(7), 549–564. https://doi.org/10.1016/j.tree.2018.04.012

Nekola, J. C., & White, P. S. (1999). The distance decay of similarity in biogeography and ecology. *Journal of Biogeography*, *26*(4), 867–878. https://doi.org/10.1046/j.1365-2699.1999.00305.x

O'Donnell, J. L., Kelly, R. P., Shelton, A. O., Samhouri, J. F., Lowell, N. C., & Williams, G. D. (2017). Spatial distribution of environmental DNA in a nearshore marine habitat. *PeerJ*, *5*, e3044.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., … Wagner, H. (2016). vegan: Community Ecology Package. R package version 2.5-2. https://CRAN.R-project.org/package=vegan

Olivier, F., San Martín, G., & Archambault, P. (2013). A new species of Streptospinigera Kudenov, 1983 (Polychaeta, Syllidae, Anoplosyllinae) from the Arctic and north-western Atlantic with a key to all species of the genus. *Polar Biology*, *36*(10), 1499–1507. https://doi.org/10.1007/s00300-013-1369-6

PAME (2016). *Hudson Bay Complex LME [Fact sheet]*. Retrieved from https://www.pame.is/index.php/document-library/ecosystem-approach-to-management-documents/large-marine-ecosystems/402-17-hudson-bay-lme/file

Payne, M. C., Reusser, D. A., & Lee, H. II. (2012). *Moderate-resolution sea surface temperature data and seasonal pattern analysis for the Arctic Ocean ecoregions: U.S. Geological Survey Open-File Report 2011–1246* (20 p.). Retrieved from http://pubs.usgs.gov/of/2011/1246/

Piepenburg, D., Archambault, P., Ambrose, W. G., Blanchard, A. L., Bluhm, B. A., Carroll, M. L., … Włodarska-Kowalczuk, M. (2011). Towards a pan-Arctic inventory of the species diversity of the macro- and megabenthic fauna of the Arctic shelf seas. *Marine Biodiversity*, *41*(1), 51–70. https://doi.org/10.1007/s12526-010-0059-7

Port, J. A., O'Donnell, J. L., Romero-Maraccini, O. C., Leary, P. R., Litvin, S. Y., Nickols, K. J., … Kelly, R. P. (2016). Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. *Molecular Ecology*, *25*(2), 527–541. https://doi.org/10.1111/mec.13481

Preez, C. D., Curtis, J. M. R., & Clarke, M. E. (2016). The structure and distribution of benthic communities on a shallow seamount (Cobb Seamount, Northeast Pacific Ocean). *PLoS ONE*, *11*(10), 1–29. https://doi.org/10.1371/journal.pone.0165513

Prinsenberg, S. J. (1984). Freshwater contents and heat budgets of James Bay and Hudson Bay. *Continental Shelf Research*, *3*(2), 191–200. https://doi.org/10.1016/0278-4343(84)90007-4

R Core Team. (2017). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, *4*, e2584. https://doi.org/10.7717/peerj.2584

Roussel, J., Paillisson, J., Tréguier, A., & Petit, E. (2015). The downside of eDNA as a survey tool in water bodies. *Journal of Applied Ecology*, *52*, 823–826. https://doi.org/10.1111/1365-2664.12428

Roy, V., Iken, K., & Archambault, P. (2015). Regional variability of megabenthic community structure across the canadian arctic. *Arctic*, *68*(2), 180–192. https://doi.org/10.14430/arctic4486

Ruiz, G. M., Carlton, J. T., Grosholz, E. D., & Hines, A. H. (1997). Global invasions of marine and estuarine habitats by non-indigenous species: Mechanisms, extent, and consequences. *American Zoologist*, *37*(6), 621–632. https://doi.org/10.1093/icb/37.6.621

Secretariat of the Convention on Biological Diversity. (2014). Global Biodiversity Outlook 4 SCBD, Convention on Biological Diversity.

Shaw, J. L. A., Clarke, L. J., Wedderburn, S. D., Barnes, T. C., Weyrich, L. S., & Cooper, A. (2016). Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. *BIOC*, *197*, 131–138. https://doi.org/10.1016/j.biocon.2016.03.010

Shokralla, S., Porter, T. M., Gibson, J. F., Dobosz, R., Janzen, D. H., Hallwachs, W., … Hajibabaei, M. (2015). Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform. *Scientific Reports*, *5*, 9687. https://doi.org/10.1038/srep09687

Socolar, J. B., Gilroy, J. J., Kunin, W. E., & Edwards, D. P. (2015). How should beta-diversity inform biodiversity conservation? *Trends in Ecology & Evolution*, *31*(1), 67–80. https://doi.org/10.1016/j.tree.2015.11.005

Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdaña, Z. A., Finlayson, M., … Robertson, J. (2007). Marine ecoregions of the World: A bioregionalization of coastal and shelf areas. *BioScience*, *57*(7), 573. https://doi.org/10.1641/B570707

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M. D. M., Breiner, H.-W., & Richards, T. A. (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Molecular Ecology*, *19*(Suppl. 1), 21–31. https://doi.org/10.1111/j.1365-294X.2009.04480.x

Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). *Environmental DNA for biodiversity and monitoring*. Oxford, UK: Oxford University Press.

Thiel, M., & Norenburg, J. (2009). *Nemertea – Ribbon worms* (pp. 369–380). Bedim Publications. Retrieved from http://www.bedim.cl/publications/Nemertea-eng-MarineBenthicFauna2009.pdf

Thomsen, P. F., Kielgast, J., Iversen, L. L., Møller, P. R., Rasmussen, M., & Willerslev, E. (2012). Detection of a diverse marine fish fauna using environmental DNA from seawater samples. *PLoS ONE*, *7*(8), 1–9. https://doi.org/10.1371/journal.pone.0041732

Thomsen, P. F., Møller, P. R., Sigsgaard, E. E., Knudsen, S. W., Jørgensen, O. A., & Willerslev, E. (2016). Environmental DNA from seawater samples correlate with trawl catches of subarctic, deepwater

fishes. *PLoS ONE*, *11*(11), e0165252. https://doi.org/10.1371/journal.pone.0165252

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, *183*, 4–18. https://doi.org/10.1016/j.biocon.2014.11.019

UNEP (The United Nations Convention on Biological Diversity). (2006). *Global biodiversity outlook 2* (81 + vii pp.). Montreal. https://doi.org/10.1093/aje/kwq338

Węsławski, J. M., & Legeżyńska, J. (2002). Life cycles of some Arctic amphipods. *Polish Polar Research*, *23*(3), 253–264. https://doi.org/10.1016/j.tourman.2010.03.016

Yamamoto, S., Masuda, R., Sato, Y., Sado, T., Araki, H., Kondoh, M., … Miya, M. (2017). Environmental DNA metabarcoding reveals local fish communities in a species-rich coastal sea. *Scientific Reports*, *7*, 40368. https://doi.org/10.1038/srep40368

Young, H. S., McCauley, D. J., Galetti, M., & Dirzo, R. (2016). Patterns, causes, and consequences of anthropocene defaunation. *Annual Review of Ecology, Evolution, and Systematics*, *47*(1), 333–358. https://doi.org/10.1146/annurev-ecolsys-112414-054142

Zhang, G. K. (2017). *Calibration of the multi-gene metabarcoding approach as an efficient and accurate biomonitoring tool* (90 p.). Master Thesis. Québec, QC: McGill University.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Leduc N, Lacoursière-Roussel A, Howland KL, et al. Comparing eDNA metabarcoding and species collection for documenting Arctic metazoan biodiversity. *Environmental DNA*. 2019;1:342–358. https://doi.org/10.1002/edn3.35

Leduc N, Lacoursière‐Roussel A, Howland KL, et al. 2019. Comparing eDNA metabarcoding and species collection for documenting Arctic metazoan biodiversity. Environmental DNA 1:342–358.

Supplementary material

Data S1. eDNA amplification

DNA amplifications were performed in a one-step dual-indexed PCR approach specifically designed for Illumina instruments by the "Plate-forme d'Analyses Génomiques" (IBIS, Université Laval). The primers were tailed on the 5' end with part of the Illumina Nextera adaptors. The following adaptor sequence (regions that anneal to the flowcell and library specific barcodes) and oligonucleotide sequences were used for amplification: AATGATACGGCGACCACCGAGATCTACAC-[INDEX]-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-[Forward primers] and reverse primers CAAGCAGAAGACGGCATACGAGAT-[INDEX]-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-[reverse primers]. Please note that primers used in this work contain Illumina specific sequences protected by intellectual property (Oligonucleotide sequences © 2007-2013 Illumina, Inc. All rights reserved. Derivative works created by Illumina customers are authorized for use with Illumina instruments and products only. All other uses are strictly prohibited). Three PCR replicates were done for each sample and each primer set. In brief, the final PCR mix of each sample replicate was composed of 12.5 µl Qiagen Multiplex Mastermix, 6.5 µl diH_2O, 1.0 µl of each primer (10 µM) and 3.0 µl of DNA. For all samples, the PCR mixture underwent an initial denaturation step at 95°C for 15 minutes, followed by 35 cycles at 94°C for 30 seconds, 54°C for 90 seconds (except for the COI2 primer set, which was at 52°C) and 72°C for 60 seconds) and a final elongation at 72°C for 10 minutes. Because barcodes were different for each sample, a negative PCR control was done for each sample and primer to ensure no false positive could happen. Following the PCR, each replicate and negative PCR control were visualized on a 1.5% agarose gel electrophoresis. If no contamination was visible, the three replicates of each sample were then pooled together. Pooled samples were purified using Axygen PCR clean up kit following the manufacture's recommended protocol and libraries

were quantified in equal molar concentrations using AccuClear Ultra High Sensitivity dsDNA Quantitation Kit using the TECAN Spark 10M Reader.

Supplementary tables

Table S1. Chordata taxa present in the eDNA data set (COI and 18S primers set added together) and the appropriate action taken.

Scientific_name	Phylum	Genus	Subphylum	Action
Ammodytes_marinus	*Chordata*	*Ammodytes*	Fish	Removed
Aplousobranchia	*Chordata*		Tunicate	Remain
Ascidia	*Chordata*		Tunicate	Remain
Ascidiacea	*Chordata*		Tunicate	Remain
Ascidiacea_colonial	*Chordata*	*Ascidiacea*	Tunicate	Remain
Ascidiella	*Chordata*		Tunicate	Remain
Boltenia_echinata	*Chordata*	*Boltenia*	Tunicate	Remain
Boltenia_villosa	*Chordata*	*Boltenia*	Tunicate	Remain
Boreogadus_saida	*Chordata*	*Boreogadus*	Fish	Removed
Brama_japonica	*Chordata*	*Brama*	Fish	Removed
Catostomus_catostomus	*Chordata*	*Catostomus*	Fish	Removed
Catostomus_commersonii	*Chordata*	*Catostomus*	Fish	Removed
Chelyosoma_macleayanum	*Chordata*	*Chelyosoma*	Tunicate	Remain
Chen_canagica	*Chordata*	*Chen*	Bird	Removed
Coregonus_artedi	*Chordata*	*Coregonus*	Fish	Removed
Coregonus_clupeaformis	*Chordata*	*Coregonus*	Fish	Removed
Cottidae	*Chordata*		Fish	Removed
Delphinapterus_leucas	*Chordata*	*Delphinapterus*	Beluga	Removed
Didemnum_albidum	*Chordata*	*Didemnum*	Tunicate	Remain
Dendrodoa_grossularia	*Chordata*	*Dendrodoa*	Tunicate	Remain
Esox_lucius	*Chordata*	*Esox*	Fish	Removed
Eumesogrammus	*Chordata*	*Eumesogrammus*	Fish	Removed
Eumicrotremus	*Chordata*	*Eumicrotremus*	Fish	Removed
Fritillaria_borealis	*Chordata*	*Fritillaria*	Tunicate	Remain
Gadus_macrocephalus	*Chordata*	*Gadus*	Fish	Removed
Glossina_pallidipes	*Chordata*	*Glossina*	Fly	Removed
Gymnelus_viridis	*Chordata*	*Gymnelus*	Fish	Removed
Gymnocanthus_tricuspis	*Chordata*	*Gymnocanthus*	Fish	Removed
Halocynthia_pyriformis	*Chordata*	*Halocynthia*	Tunicate	Remain
Icelus	*Chordata*	*Icelus*	Fish	Removed
Icelus_bicornis	*Chordata*	*Icelus*	Fish	Removed
Larus_californicus	*Chordata*	*Larus*	Bird	Removed
Lateolabrax_maculatus	*Chordata*	*Lateolabrax*	Fish	Removed
Liparidae	*Chordata*		Fish	Removed

Liparis_inquilinus	*Chordata*	*Liparis*	Fish	Removed
Lota_lota	*Chordata*	*Lota*	Fish	Removed
Lumpenus_lampretaeformis	*Chordata*	*Lumpenus*	Fish	Removed
Lycodes_mucosus	*Chordata*	*Lycodes*	Fish	Removed
Maccullochella_macquariensis	*Chordata*	*Maccullochella*	Fish	Removed
Mallatus_villosus	*Chordata*	*Mallatus*	Fish	Removed
Mallotus_catervarius	*Chordata*	*Mallotus*	Fish	Removed
Mallotus_villosus	*Chordata*	*Mallotus*	Fish	Removed
Molgula	*Chordata*	*Molgula*	Tunicate	Remain
Molgula_retortiformis	*Chordata*	*Molgula*	Tunicate	Remain
Molgulidae	*Chordata*		Tunicate	Remain
Myoxocephalus	*Chordata*	*Myoxocephalus*	Fish	Removed
Myoxocephalus_quadricornis	*Chordata*	*Myoxocephalus*	Fish	Removed
Myoxocephalus_scorpius	*Chordata*	*Myoxocephalus*	Fish	Removed
Oikopleura	*Chordata*		Tunicate	Remain
Osmerus_mordax	*Chordata*	*Osmerus*	Fish	Removed
Ovis_aries	*Chordata*	*Ovis*	Sheep	Removed
Pagophilus_groenlandicus	*Chordata*	*Pagophilus*	Seal	Removed
Pelonaia_corrugata	*Chordata*	*Pelonaia*	Tunicate	Remain
Perca_flavescens	*Chordata*	*Perca*	Fish	Removed
Pisces	*Chordata*		Fish	Removed
Pungitius_pungitius	*Chordata*	*Pungitius*	Fish	Removed
Pusa_hispida	*Chordata*	*Pusa*	Seal	Removed
Rangifer_tarandus	*Chordata*	*Rangifer*	Reindeer	Removed
Rhinichthys_cataractae	*Chordata*	*Rhinichthys*	Fish	Removed
Salvelinus_fontinalis	*Chordata*	*Salvelinus*	Fish	Removed
Stichaeidae	*Chordata*		Fish	Removed
Stichaeus_punctatus	*Chordata*	*Stichaeus*	Fish	Removed
Stichaeus_punctatus_punctatus	*Chordata*	*Stichaeus*	Fish	Removed
Styela_rustica	*Chordata*	*Styela*	Tunicate	Remain
Styela_gibbsii	*Chordata*	*Styela*	Tunicate	Remain
Styelidae	*Chordata*		Tunicate	Remain
Triglops	*Chordata*		Fish	Removed

Table S2. Barque 1.5.1 specific commands for preparation and analysis of pair-end reads. The sequences from the different COI and 18S primers set were added together.

Main step of filtration	Number of eDNA reads in Churchill harbour	Number of eDNA reads in Deception Bay harbour	Number of eDNA reads in Iqaluit harbour
Number of raw forward and reverse reads	10 553 694	5 767 153	7 960 264
Remove raw reads with low quality	10 171 078	5 547 150	7 679 365
Remaining reads after merging forward and reverse reads	9 635 596	5 136 684	7 255 383
Remove reads with incorrect length	8 535 538	4 600 181	6 405 199
Remove chimeric reads	8 170 548	4 340 911	6 008 858
Number of final reads	7 262 004	3 578 713	4 880 549
Number of reads with successful BLAST at genus level	574 581	113 236	206 269
Number of reads with successful BLAST at species level	478 046	95 651	203 245

Table S3. All taxa successfully blasted at the genus level within eDNA dataset. Total number of reads obtained for COI and 18S primers combined within the three ports, within field control and laboratory extraction negative control.

Genus	Phylum	Number of all samples sequences	Number of all field control sequences	Number of all laboratory negative control sequences
Abarenicola	Annelida	4	0	0
Acartia	Arthropoda	134619	10	1
Actinauge	Cnidaria	10	0	0
Adineta	Rotifera	2	0	0
Aeginopsis	Cnidaria	12	0	0
Alcyonidioides	Bryozoa	146	11	0
Alcyonidium	Bryozoa	765	43	37
Ameronothrus	Arthropoda	199	0	0
Ammodytes	Chordata	27	0	0
Amphichaeta	Annelida	20	0	0
Amphicorina	Annelida	48	0	0
Amphiporus	Nemertea	36	0	0
Amphitrite	Annelida	1950	0	0
Amphiura	Echinodermata	6	0	0
Anoplodactylus	Arthropoda	2	0	0
Aphelochaeta	Annelida	4255	1	2
Apistobranchus	Annelida	10	1	0
Arcteonais	Annelida	133	0	0
Arenicola	Annelida	486	0	0
Argis	Arthropoda	3	0	0
Aricidea	Annelida	29	0	0
Ascomorpha	Rotifera	71	0	0
Aulactinia	Cnidaria	45	0	0
Aurelia	Cnidaria	574	233	0
Aurospio	Annelida	151	0	1
Axionice	Annelida	53	0	0
Balanus	Arthropoda	104150	9	0
Beroe	Cyclocoela	221	3	0
Bipalponephtys	Annelida	29	0	0
Boltenia	Chordata	51	0	0
Bonneviella	Hydrozoa	2	0	0
Boreogadus	Chordata	1613	0	0
Bosmina	Arthropoda	8	0	0
Bougainvillia	Cnidaria	324	0	0
Brama	Chordata	2	0	0
Bylgides	Annelida	288	0	0
Calanus	Arthropoda	30	3	0
Calycella	Cnidaria	70	1	0

Candona	Arthropoda	77	0	0
Capitella	Annelida	7	0	0
Caprella	Arthropoda	9	0	0
Cardites	Mollusca	16	0	0
Castrella	Platyhelminthes	2	0	0
Catablema	Cnidaria	128	0	0
Catostomus	Chordata	57	0	0
Cauloramphus	Bryozoa	106	0	0
Celleporella	Bryozoa	26	0	0
Centropages	Arthropoda	1710	0	0
Cephalodella	Rotifera	178	0	0
Cephalothrix	Nemertea	3104	0	0
Cerebratulus	Nemertea	14	8	0
Cernosvitoviella	Annelida	2	0	0
Chaetogaster	Annelida	2049	0	0
Chaetonotus	Gastrotricha	13	0	0
Chaetozone	Annelida	1195	0	0
Chelyosoma	Chordata	14	0	0
Chen	Chordata	3	0	0
Chitinopoma	Annelida	111	0	0
Chlamys	Mollusca	50	0	0
Chone	Annelida	48	0	0
Chydorus	Arthropoda	5	0	0
Cirratulus	Annelida	23	0	1
Cirrophorus	Annelida	14	0	0
Clathrina	Calcinea	18	0	0
Clavactinia	Hydrozoa	6	0	0
Clione	Mollusca	1641	1	1
Clymenura	Annelida	52	0	0
Corbula	Mollusca	7	0	0
Coregonus	Chordata	121	0	0
Craniella	Spirophorida	38	0	0
Cucumaria	Echinodermata	142	0	0
Cyanea	Cnidaria	935	0	0
Cyclopina	Arthropoda	9	0	0
Cypridopsis	Arthropoda	14	0	0
Cypris	Arthropoda	14	0	0
Cytheromorpha	Arthropoda	6	0	0
Daphnia	Arthropoda	4	0	0
Delphinapterus	Chordata	587	0	0
Dendrodoa	Chordata	63	0	0
Dermatophagoides	Arthropoda	3	0	0
Desoria	Arthropoda	41	35	0
Earleria	Cnidaria	115	0	0
Echinogammarus	Rotifera	4	0	0
Echiurus	Annelida	280	0	0

Ectopleura	Cnidaria	5	0	0
Enipo	Annelida	105	0	0
Ennucula	Mollusca	2761	0	0
Ephydatia	Porifera	39	0	0
Erinaceusyllis	Annelida	15	0	0
Erpobdella	Annelida	11	0	0
Esox	Chordata	4	0	0
Eudendrium	Hydrozoa	76	0	0
Eulalia	Annelida	11	0	0
Eumicrotremus	Chordata	5	0	0
Eunapius	Porifera	129	0	0
Eunoe	Annelida	174	0	0
Euphysa	Cnidaria	233	0	171
Eurycercus	Arthropoda	8	0	0
Eurytemora	Arthropoda	10976	443	0
Flabelligera	Annelida	16	0	0
Flabellina	Mollusca	32	0	0
Florometra	Echinodermata	10	0	0
Gadus	Chordata	7	0	0
Galathowenia	Annelida	4950	0	0
Gammaracanthus	Arthropoda	2	0	0
Gammarus	Arthropoda	659	73	15
Gattyana	Annelida	81	0	0
Gersemia	Cnidaria	31	0	0
Glossina	Chordata	6	0	3
Glycera	Annelida	23	0	0
Golfingia	Sipuncula	56	0	0
Gonothyraea	Cnidaria	67	0	0
Gorgonocephalus	Echinodermata	3	0	0
Gymnelus	Chordata	2	0	0
Gymnocanthus	Chordata	19	0	0
Gyraulus	Mollusca	47	0	0
Halcampoides	Anthozoa	11	0	1
Halecium	Hydrozoa	5	0	0
Halicephalobus	Nematoda	71	70	0
Halichaetonotus	Gastrotricha	26	24	0
Halichondria	Porifera	814	1	0
Haliclona	Porifera	69	0	0
Halisarca	Porifera	422	0	0
Halocynthia	Chordata	511	0	0
Hamigera	Poecilosclerida	8	0	0
Harmothoe	Annelida	7292	0	0
Heterolepidoderma	Gastrotricha	2	0	0
Hiatella	Mollusca	13904	0	2
Holopedium	Arthropoda	33	0	3
Homarus	Arthropoda	8	0	8

Hyalinella	Bryozoa	2	0	0
Hyas	Arthropoda	10	0	0
Hybocodon	Hydrozoa	6	0	0
Hybomitra	Arthropoda	5	0	0
Hydra	Cnidaria	232	0	0
Hydropsyche	Arthropoda	206	0	0
Hymeniacidon	Hadromerida	283	0	0
Hyperia	Arthropoda	2	0	0
Icelus	Chordata	6	0	0
Iophon	Poecilosclerida	115	0	0
Keratella	Rotifera	653	0	0
Lampocteis	Cyclocoela	40	1	1
Laomedea	Hydrozoa	149	0	0
Larochella	Mollusca	672	36	570
Larus	Chordata	27	0	0
Lateolabrax	Chordata	14	13	0
Lebbeus	Arthropoda	19	0	0
Leiosolenus	Mollusca	106	0	0
Leptasterias	Echinodermata	54	0	0
Leptocythere	Arthropoda	4	0	1
Leptodiaptomus	Arthropoda	5	0	0
Leuconia	Calcaronea	124	0	0
Leucothea	Ctenophora	106	0	3
Limacina	Mollusca	16	0	0
Limnodrilus	Annelida	25	0	0
Lineus	Nemertea	5	0	0
Liparis	Chordata	6	0	0
Littorina	Mollusca	3708	4	0
Lophaster	Echinodermata	15	0	0
Lota	Chordata	28	0	0
Loxosomella	Entoprocta	4	0	0
Lucernaria	Cnidaria	12	0	0
Lumbricillus	Annelida	221	65	0
Lumbriculus	Annelida	19	0	0
Lumbricus	Annelida	2	0	0
Lumbrineris	Annelida	136	0	0
Lumpenus	Chordata	53	0	0
Lycodes	Chordata	8	0	0
Lymnaea	Mollusca	7	0	0
Maccullochella	Chordata	10	0	0
Macoma	Mollusca	804	0	0
Macrocyclops	Arthropoda	34	34	0
Macrophiothrix	Echinodermata	40	1	0
Maldane	Annelida	11	0	0
Mallotus	Chordata	5	0	0
Margarites	Mollusca	28	0	0

Melicertum	Cnidaria	28	0	0
Mertensia	Typhlocoela	12204	9	0
Mesenchytraeus	Annelida	2	0	0
Mesochra	Arthropoda	9	0	0
Metridia	Arthropoda	3	0	0
Microsetella	Arthropoda	684	1	0
Microstomum	Platyhelminthes	19	0	0
Mideopsis	Arthropoda	2	0	0
Mitrocomella	Cnidaria	177	0	0
Mopalia	Mollusca	370	0	0
Motobdella	Annelida	175	0	0
Myoxocephalus	Chordata	34	0	0
Mysis	Arthropoda	394	160	0
Mytilus	Mollusca	1388	30	13
Myxilla	Poecilosclerida	13	0	0
Nais	Annelida	2071	1	0
Neoleptophlebia	Arthropoda	2	0	0
Nephtys	Annelida	3067	0	1
Nereis	Annelida	2974	1	0
Nicolea	Annelida	5	0	0
Nicomache	Annelida	17	0	0
Notholca	Rotifera	45	0	0
Notommata	Rotifera	4	0	0
Obelia	Cnidaria	20	0	0
Ochlerotatus	Arthropoda	35	25	0
Oithona	Arthropoda	20332	0	2
Onisimus	Arthropoda	4	0	0
Opercularella	Cnidaria	824	1	0
Ophelia	Annelida	1018	0	0
Ophelina	Annelida	774	0	0
Ophiopholis	Echinodermata	23455	0	0
Ophiura	Echinodermata	10321	1	0
Orthopyxis	Cnidaria	51	0	0
Ovis	Chordata	19	19	0
Owenia	Annelida	8	0	0
Pachypellina	Porifera	2	0	0
Pagophilus	Chordata	18	0	0
Palio	Mollusca	6	0	0
Palliolum	Mollusca	10	0	0
Panopea	Mollusca	2	0	0
Paranais	Annelida	12	0	0
Paranerilla	Annelida	500	0	0
Peachia	Anthozoa	10	0	0
Pectinaria	Annelida	148118	3	35
Pelonaia	Chordata	284	0	0
Pentamera	Echinodermata	17	0	0

Perca	Chordata	4	0	0
Pholoe	Annelida	1242	0	0
Phyllodoce	Annelida	57109	5	2
Physa	Mollusca	44	0	0
Physella	Mollusca	79	0	0
Pisidium	Mollusca	22	0	0
Pista	Annelida	57	0	0
Polyarthra	Rotifera	969	0	1
Polycirrus	Annelida	6319	0	0
Polydora	Annelida	5	0	0
Praxillella	Annelida	581	0	0
Priapulopsis	Scalidophora	21	0	0
Priapulus	Priapulida	9	0	0
Prionospio	Annelida	80	0	0
Pristina	Unknown	3	0	0
Provortex	Platyhelminthes	29	0	0
Pseudocalanus	Arthropoda	256076	33772	53
Pseudoscalibregma	Annelida	153	0	0
Pseudosuberites	Porifera	190	0	0
Psolus	Echinodermata	215	0	0
Pteraster	Echinodermata	4	0	0
Pterocirrus	Annelida	41	0	0
Pungitius	Chordata	28	0	0
Pusa	Chordata	3	0	0
Pyganodon	Mollusca	15	0	0
Pygospio	Annelida	4	0	0
Rangifer	Chordata	8	0	0
Rhinichthys	Chordata	43	0	0
Rhizoglyphus	Arthropoda	2	0	0
Rhysotritia	Arthropoda	387	38	12
Rotaria	Rotifera	5	0	0
Salvelinus	Chordata	18	0	0
Sarsia	Cnidaria	40	0	0
Scalibregma	Annelida	375	4	1
Schuchertinia	Hydrozoa	35	0	0
Scoletoma	Annelida	875	0	0
Semibalanus	Arthropoda	13	0	0
Sertularella	Cnidaria	11	0	0
Sida	Arthropoda	194	0	0
Simocephalus	Arthropoda	28	0	0
Solaster	Echinodermata	15	0	0
Solmundella	Hydrozoa	506	0	0
Specaria	Annelida	166	0	0
Sperchon	Arthropoda	3	0	0
Sphaerium	Mollusca	113	0	0
Spio	Annelida	661	0	0

Spongilla	Porifera	119	0	0
Stagnicola	Mollusca	90	0	0
Stegophiura	Echinodermata	12	0	0
Stenostomum	Platyhelminthes	23	0	0
Stichaeus	Chordata	3	0	0
Strongylocentrotus	Echinodermata	5209	0	0
Styela	Chordata	12	0	0
Stylaria	Annelida	87	0	0
Suberites	Porifera	8	0	0
Synchaeta	Rotifera	15	0	5
Synmerosyllis	Annelida	7	0	0
Synute	Calcaronea	17	0	0
Tectura	Mollusca	1836	0	0
Terebellides	Annelida	12	0	0
Tetrastemma	Nemertea	19	0	0
Thelepus	Annelida	6862	1	0
Thyonidium	Echinodermata	196	0	0
Thysanoessa	Arthropoda	1163	0	0
Tiaropsis	Hydrozoa	2	0	0
Tisbe	Arthropoda	68	6	2
Tonicella	Mollusca	5702	0	2
Vulgarogamasus	Arthropoda	14	0	0
Zaus	Arthropoda	999	18	154

Table S4. Genera found in the negative field controls and the appropriate action taken against the contamination according to COI and 18S primers set.

Genus	Number of all field control sequences	Number of all sample sequences	Percentage of contamination (%)	Contamination approaches	Primers set	Type of animal/distribution
Alcyonidioides	11	158	7	removed	18S	Bryozoan - marine
Alcyonidium	43	587	7	removed	18S	Bryozoan - marine
Apistobranchus	1	9	11	removed	18S	Worms - arctic marine waters
Aurelia	232	295	79	removed	18S	Jellyfish - marine
Calanus	3	24	123	removed	18S	Zooplankton - marine
Eurytemora	443	5	8860	removed	18S	Zooplankton - marine, brackish, fresh waters
Halicephalobus	70	1	7000	removed	18S	Nematode - parasite
Halichaetonotus	24	2	1200	removed	18S	Gastrotrich - marine
Lampocteis	1	36	3	removed	18S	Comb jelly - subtropical marine waters
Larochella	36	49	734	removed	18S	Gastropod - marine
Lumbricillus	65	47	138	removed	18S	Worms - marine
Mysis	159	157	101	removed	18S	Crustacean - marine
Mytilus	25	516	5	removed	18S	Bivalve mollusks - marine
Pseudocalanus	33696	87731	38	remain	18S	Zooplankton - arctic marine waters
Scalibregma	2	72	3	removed	18S	Worms - arctic marine waters
Zaus	18	734	3	removed	18S	Crustacean - marine
Aphelochaeta	1	4226	0.02	remain	18S	Worms - marine
Boroe	3	216	1	remain	18S	Comb jelly - marine
Calycella	1	71	1	remain	18S	Cnidarian -marine
Halichondria	1	705	0.1	remain	18S	Sponge - marine
Mertensia	9	12352	0.1	remain	18S	Comb jelly - marine
Nereis	1	318	0.3	remain	18S	Worms - marine
Opercularella	1	806	0.1	remain	18S	Cnidaria - marine
Pectinaria	1	30881	0.001	remain	18S	Worms - marine
Phyllodoce	2	21422	0.01	remain	18S	Worms - marine
Aurelia	1	41	2	removed	COI	Jellyfish - marine
Cerebratulus	8	6	133	removed	COI	Worms -arctic marine waters
Gammarus	68	518	13	removed	COI	Crustacean - marine
Macrocyclops	34	0		removed	COI	Zooplankton - freshwater
Macrophiothrix	1	39	3	removed	COI	Brittle star - marine
Tisbe	6	58	10	removed	COI	Crustacean - marine

Acartia	10	134600	0.01	remain	COI	Zooplankton - marine
Balanus	9	104141	0.01	remain	COI	Crustacean - marine
Clione	1	1638	0.1	remain	COI	Sea angel - cold marine waters
Littorina	4	3693	0.1	remain	COI	Sea snails - marine
Microsetella	1	683	0.1	remain	COI	Zooplankton - marine
Mysis	1	80	1	remain	COI	Crustacean - marine
Mytilus	5	754	0.7	remain	COI	Bivalve mollusks - marine
Nais	1	1839	0.1	remain	COI	Worms - brackish and freshwater
Ophiura	1	10314	0.01	remain	COI	Brittle star - marine
Pectinaria	1	117866	0.001	remain	COI	Worms - marine
Phyllodoce	3	35675	0.01	remain	COI	Worms - marine
Pseudocalanus	70	131765	0.1	remain	COI	Zooplankton - arctic marine waters
Scalibregma	2	296	0.7	remain	COI	Worms - arctic marine waters
Thelepus	1	6861	0.01	remain	COI	Worms - arctic marine waters

Table S5. Summary of PERMANOVA statistics tests on marine invertebrates communities for the phylum relative abundance (number of taxa), Pielou evenness index and alpha richness. The analyses were performed with method = "bray" for phylum relative abundance while it was performed with method = "euclidian" for Pielou evenness and alpha richness.

Evaluated parameter	Source of variation	PERMANOVA		
		F-value	R2	Pr (> f)
eDNA community composition	Harbour	40.177	0.416	< 0.001
	Churchill vs. Iqaluit	46.785	0.384	0.003
	Churchill vs. Deception Bay	34.180	0.313	0.003
	Iqaluit vs. Deception Bay	38.371	0.335	0.003
Species collection community composition	Harbour	6.706	0.078	< 0.001
	Churchill vs. Iqaluit	8.258	0.067	0.003
	Churchill vs. Deception Bay	6.780	0.056	0.003
	Iqaluit vs. Deception Bay	4.449	0.048	0.015
Phylum relative abundance	Method	43.708	0.337	< 0.001
Churchill	Sp. collection vs eDNA	12.243	0.265	< 0.001
Deception Bay	Sp. collection vs eDNA	38.305	0.615	< 0.001
Iqaluit	Sp. collection vs eDNA	64.523	0.723	< 0.001
eDNA Pielou evenness index	Harbour	10.663	0.372	< 0.001
	Churchill vs. Iqaluit	5.55	0.19	0.09
	Churchill vs. Deception Bay	5.356	0.182	0.087
	Iqaluit vs. Deception Bay	20.431	0.460	0.003
Species collection Pielou evenness index	Harbour	1.9	0.08	0.2
Alpha richness	Method			
Churchill	Sp. collection vs eDNA	77.471	0.695	< 0.001
Deception Bay	Sp. collection vs eDNA	0.868	0.035	0.365
Iqaluit	Sp. collection vs eDNA	53.641	0.691	< 0.001

Table S6. Summary of the correlation between dissimilarity and distance across the sites within Churchill, Deception Bay and Iqaluit harbours based on incidence data for the different sampling methods.

Method	Harbour	Correlation test	Simulated p-value	R^2
eDNA				
	Churchill	Cor.test	< 0.001	0.13
	Deception Bay	Mantel.rtest	0.01	0.23
	Iqaluit	Mantel.rtest	0.02	0.094
Trawl, grab, cores				
	Churchill	Cor.test	0.5	0.004
	Deception Bay	Mantel.rtest	0.02	0.12
	Iqaluit	Mantel.rtest	0.01	0.14
Net tows				
	Churchill	Cor.test	0.2	0.014
	Deception Bay	Mantel.rtest	0.01	0.26
	Iqaluit	Mantel.rtest	1.0	0.16

Table S7. Summary of the main phyla identified by sampling collection methods among Churchill, Deception Bay and Iqaluit harbours and their respective presence in BOLD and SILVA public genetic databases.

Phylum	Number of organisms identified with trad. methods	Number of identified organisms present in BOLD database	Number of identified organisms present in SILVA database	Total number of identified organisms present in genetic database (BOLD or SILVA)	Percentage of the identified organisms collected with trad. methods available in genetic database (BOLD or SILVA)
Annelida	140	80	65	95	67.9
Arthropoda	167	89	43	94	56.3
Brachiopoda	1	0	1	1	100.0
Bryozoa	42	5	10	12	28.6
Cephalorhyncha	2	0	0	0	0.0
Chordata	19	13	6	14	73.7
Cnidaria	16	9	0	9	56.3
Echinodermata	9	7	2	7	77.8
Mollusca	77	37	28	42	54.5
Myzozoa	1	0	0	0	0.0
Porifera	2	0	0	0	0.0

Supplementary figures

Figure S1. The number of Operational taxonomic units (OTUs) assigned and not assigned on NCBI to the genus level for the COI primers set (COI1 and COI2) and the 18S primers set (18S1 and 18S2). The assigned OTUs are represented by the grey section of the barplot while the non assigned OTUs are represented by the red section of the barplot. The percentages in red represented the mean of the 4 primers of unassigned OTUs in each port.

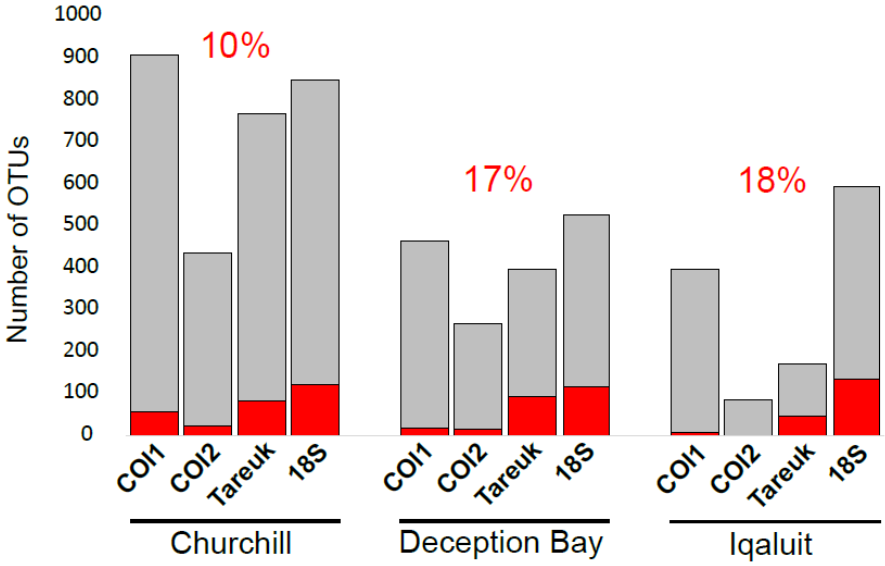


Figure S2. Individual-based rarefaction curves of eDNA, benthos and zooplankton genera for Churchill (blue), Deception Bay (yellow) and Iqaluit harbours (magenta) based on incidence data.

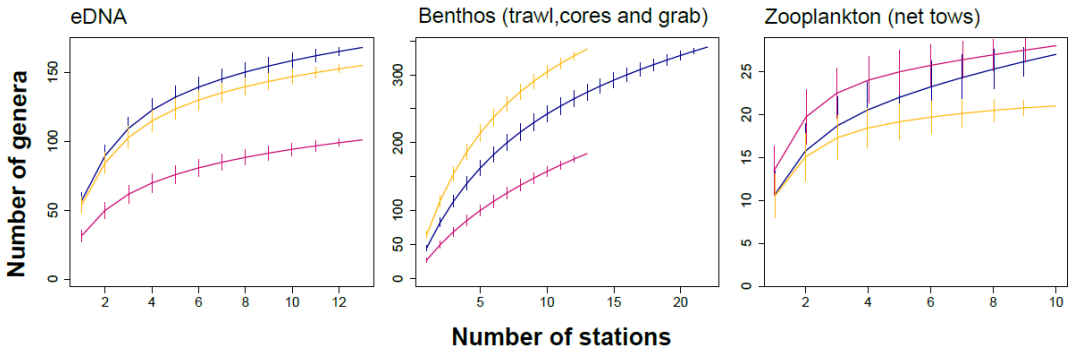
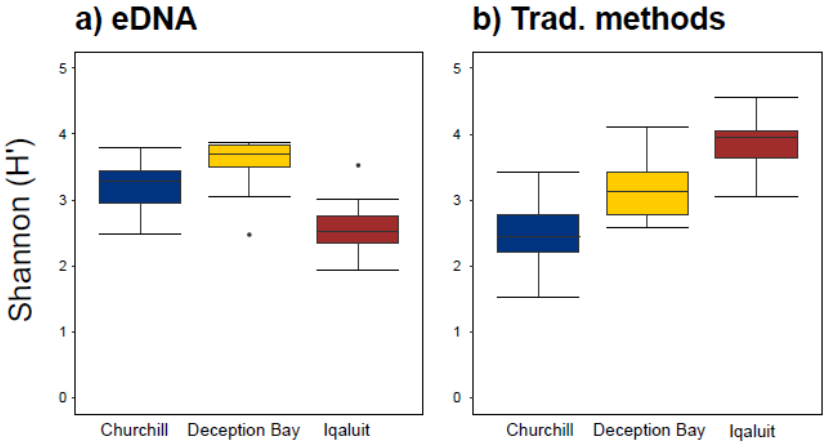


Figure S3. Boxplot on alpha diversity for Shannon biodiversity index in Churchill, Deception Bay and Iqaluit harbours. These analyses were performed on abundance data with Hellinger transformation, COI and 18S primer sets are added together for the eDNA boxplot.



Received: 3 July 2019 | Revised: 3 September 2019 | Accepted: 4 September 2019

DOI: 10.1002/ece3.5676

ORIGINAL RESEARCH

Ecology and Evolution Open Access WILEY

Evidence for host effect on the intestinal microbiota of whitefish (*Coregonus* sp.) species pairs and their hybrids

Maelle Sevellec | Martin Laporte | Alex Bernatchez | Nicolas Derome | Louis Bernatchez

Institut de Biologie Intégrative et des Systèmes (IBIS), Pavillon Charles-Eugène-Marchand, Université Laval, Québec, QC, Canada

Correspondence
Maelle Sevellec, Institut de Biologie Intégrative et des Systèmes (IBIS), Pavillon Charles-Eugène-Marchand, 1030, Avenue de la Médecine, Université Laval, Québec, QC G1V 0A6, Canada.
Email: maelle.sevellec@gmail.com

Funding information
Natural Sciences and Engineering Research Council of Canada; Québec-Océan

Abstract

Investigating relationships between microbiota and their host is essential toward a full understanding of how animal adapt to their environment. Lake Whitefish offers a powerful system to investigate processes of adaptive divergence where the dwarf, limnetic species evolved repeatedly from the normal, benthic species. We compared the transient intestinal microbiota between both species from the wild and in controlled conditions, including their reciprocal hybrids. We sequenced the 16s rRNA gene V3-V4 regions to (a) test for parallelism in the transient intestinal microbiota among sympatric pairs, (b) test for transient intestinal microbiota differences among dwarf, normal, and hybrids reared under identical conditions, and (c) compare intestinal microbiota between wild and captive whitefish. A significant host effect on microbiota taxonomic composition was observed when all lakes were analyzed together and in three of the five species pairs. In captive whitefish, host effect was also significant. Microbiota of both reciprocal hybrids fell outside of that observed in the parental forms. Six genera formed a bacterial core which was present in captive and wild whitefish, suggesting a horizontal microbiota transmission. Altogether, our results complex interactions among the host, the microbiota, and the environment, and we propose that these interactions define three distinct evolutionary paths of the intestinal microbiota.

KEYWORDS
captive whitefish, intestinal microbiota, speciation, wild whitefish

1 | INTRODUCTION

Woese (1998) referred the Earth as a microbial planet, where macro-organisms are recent additions. Indeed, an increasing number of studies have highlighted the substantial impact of microbiota on their host genes (Hooper et al., 2001; Rawls, Samuel, & Gordon, 2004) and that microbiota may be transmitted across generations in both animals and plants (Rosenberg & Zilber, 2016). In fishes in particular, the mother allocates antimicrobial compounds to the eggs before spawning (Hanif, Bakopoulos, & Dimitriadis, 2004; Wilkins, Rogivue, Fumagalli, & Wedekind, 2015). This maternal selection of bacteria influences the first bacteria that will be in contact with the sterile larvae during hatching (Llewellyn, Boutin, Hoseinifar, & Derome, 2014). Clearly then, a holistic understanding of macro-organisms biodiversity requires the investigation of their association with microbiota and their co-evolution (Miller, Svanbäck, & Bohannan, 2018).

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2019 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

The hologenome concept stipulates that the genome of the host and the microbiome (i.e., sum of the genetic information of the microbiota) act in consortium as a unique biological entity, that is, the holobiont (Rosenberg & Zilber, 2013). Consequently, the microbiota may be involved in host reproductive isolation, either in the form of a prezygotic barrier by influencing the host's mate choice by chemosensory signals (Brucker & Bordenstein, 2012; Damodaram, Ayyasamy, & Kempraj, 2016; Shropshire & Bordenstein, 2016) or in the form of a postzygotic barrier by producing genome and microbiome incompatibilities in hybrids, in accordance with the Bateson, Dobzhansky, and Muller model of genetic incompatibilities (Brucker & Bordenstein, 2012; Dobzhansky, 1937; Muller, 1942). Because the bacterial community of the gastrointestinal tract is implicated in many critical functions essential for development and immune responses, the intestinal microbiota could play an important role on its host's adaptive potential (Alberdi, Aizpurua, Bohmann, Zepeda-Mendoza, & Gilbert, 2016; Macke, Tasiemski, Massol, Callens, & Decaestecker, 2017; Rosenberg & Zilber, 2013).

Fishes as a group comprise the greatest taxonomic diversity of vertebrates and a major food resource for human populations (Béné et al., 2015; Nelson, 2006), yet little is known about the relationship with their microbiota compared with the already well-characterized mammals and insect microbiota (Clements, Angert, Montgomery, & Choat, 2014). The Lake Whitefish (*Coregonus clupeaformis*) is a well-studied system that represents a continuum in the early stage of speciation where sympatric species pairs of dwarf and normal species evolved independently in several lakes in northeastern North America (Bernatchez et al., 2010; Rougeux, Bernatchez, & Gagnaire, 2017). The normal species is specialized for using the trophic benthic niche, feeding on diverse prey as zoobenthos and molluscs. It is characterized by rapid growth, late sexual maturity, and a long lifespan (Bodaly, 1979; Landry & Bernatchez, 2010). In contrast, the dwarf whitefish is a limnetic specialist which feeds almost exclusively on zooplankton and is characterized by slower growth, early sexual maturation, and shorter lifespan compared with the normal species. Previous transcriptomic studies revealed overexpression of genes implicated with survival functions (e.g., enhanced swimming performance for predator avoidance, detoxification) in dwarf whitefish, whereas normal whitefish show overexpression of genes associated with growth functions (Bernatchez et al., 2010; StCyr, Derome, & Bernatchez, 2008). Moreover, many other physiological, morphological, and behavioral traits display parallel differences among these two whitefish species that correspond to their respective trophic specialization (Bernatchez et al., 2010; Dalziel, Laporte, Guderley, & Bernatchez, 2017; Dalziel, Laporte, Rougeux, Guderley, & Bernatchez, 2016; Dalziel, Martin, Laporte, Guderley, & Bernatchez, 2015; Gagnaire, Normandeau, Pavey, & Bernatchez, 2013; Jeukens, Bittner, Knudsen, & Bernatchez, 2009; Laporte, Dalziel, Martin, & Bernatchez, 2016; Laporte et al., 2015). Thus, the recent speciation and the clear trophic segregation make the whitefish species pair an excellent model to study the role of intestinal microbiota in the context of ecological speciation.

Two previous studies documented the variation in two microbial niches in Lake Whitefish species pairs: the kidney and the intestinal adherent communities (Sevellec, Derome, & Bernatchez, 2018; Sevellec et al., 2014). Although we observed parallel patterns of differentiation between normal and dwarf species in the bacterial kidney communities, no clear evidence for parallelism was observed in the adherent intestinal microbiota. However, the water bacterial community was distinct from the adherent intestinal microbiota, suggesting an intrinsic properties of the host microbiota (Sevellec et al., 2018). There is increasing evidence that allochthonous microbial communities (hereafter the transient microbiota) ingested from the environment by the host play a significant role in the overall gut microbiota, either by stimulating colonization resistance or by providing additional functions to the host (e.g., David et al., 2014). However, few studies have tested for parallelism patterns in fish intestinal microbiota (Baldo et al., 2017; Baldo, Riera, Tooming-Klunderud, Albà, & Salzburger, 2015; Hata et al., 2014; Sevellec et al., 2014; Smith, Snowberg, Caporaso, Knight, & Bolnick, 2015; Sullam et al., 2015). Also, the effect of the hybridization of two recently diverged species on microbiota composition is still poorly documented (Guivier et al., 2017).

The main goal of this study was to document the transient intestinal microbiota taxonomic composition of Lake Whitefish species pairs and their hybrids in natural and controlled environment. We investigated the transient intestinal microbiota in five wild species pairs of whitefish to estimate the within- and between-lake variation and tested for parallelism among lakes. Secondly, we characterized the taxonomic composition of transient intestinal microbiota on dwarf, normal, and first-generation hybrids reared in common garden in order to test the influence of the whitefish host on the microbiota in the same controlled conditions and under two different diets.

2 | MATERIALS AND METHODS

2.1 | Sample collection of wild whitefish

Lake Whitefish were sampled from May to July 2013 in Cliff, Indian, and Webster lakes in Maine, United States, and in East and Témiscouata lakes in Québec, Canada (Table 1). Fish were dissected in the field in sterile conditions as detailed previously (Sevellec et al., 2018). The intestine was cut at the hindgut end level (posterior part of the intestine), and the digesta were aseptically squeezed to collect the alimentary bolus. All samples of alimentary bolus were transported to the laboratory and kept at −80°C until further processing.

2.2 | Experimental crosses, rearing conditions, and sample collection for captive whitefish

In November 2013, 32 fish representing four cross types, dwarf (D♀×D♂), normal (N♀×N♂), and their reciprocal hybrids (F1 D♀×N♂ and F1 N♀×D♂), were pooled together in three tanks (eight fish/form/tank) (Figure 1). Experimental cross design was as described previously (Dalziel et al., 2015; Laporte et al., 2016). The protocol

TABLE 1 Number and locations of samples, sampling dates for each captive and wild whitefish populations or group

Origin	Form	Sample size	Sampling date	Coordinates
Cliff	DD	12	13 to 14 June 2013	46°23′59″N, 69°15′11″W
	NN	12		
East	DD	10	2 to 4 July 2013	47°11′15″N, 69°33′41″W
	NN	13		
Indian	DD	12	10 to 11 June 2013	46°15′32″N, 69°17′29″W
	NN	13		
Témiscouata	DD	10	28 to 30 May 2013	47°40′04″N, 68°49′03″W
	NN	14		
Webster	DD	3	12 to 13 June 2013	46°09′23″N, 69°04′52″W
	NN	12		
Common Garden 1	DD	7	12 November 2013 to 09 June 2014	LARSA
	NN	5		
	DH	7		
	NH	6		
Common Garden 2	DD	5	12 November 2013 to 10 June 2014	LARSA
	NN	4		
	DH	6		
	NH	6		
Common Garden 3	DD	8	12 November 2013 to 11 June 2014	LARSA
	NN	6		
	DH	6		
	NH	8		

Abbreviations: DD, dwarf whitefish; DH, hybrid F1 D♀×N♂; NH, hybrid F1 N♀×D♂; NN, normal whitefish.

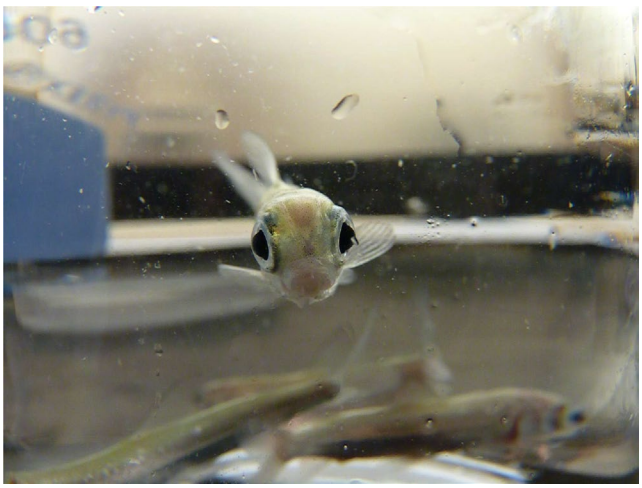


FIGURE 1 Picture of a juvenile captive hybrid whitefish at the beginning of experiment (November 2013)

used for whitefish eggs fertilization and creating the parental generation is detailed in Appendix S1. Fish were separated in three tanks sharing the same experimental conditions (water, food, pH, and temperature) for seven months. Juvenile whitefish were fed on two types of food: *Artemia* and dry food pellet BioBrood (Bio-Oregon®) (Flüchter, 1982; Zitzow & Millard, 1988). Fish were reassigned to their group of origin based on genetic allocation using mitochondrial and microsatellite DNA markers (Appendix S1). In June 2014, fish were euthanized with MS-222 and dissected immediately in sterile conditions, as described previously (Sevellec et al., 2018). Samples were kept at 80°C until further processing. This study was approved under Institutional Animal Care and Use Committee protocol 2008-0106 at Laval University.

2.3 | Whitefish microbiota: DNA extraction, amplification, and sequencing

The alimentary boluses of all fish were extracted using a modification of the QIAmp© Fast DNA stool mini kit (QIAGEN) (Appendix S1). In order to construct the community library, a region ~250 bp in the 16S rRNA gene, covering the V3-V4 region, was amplified (detailed in Appendix S1) using specific primers with Illumina barcoded adapters Bakt_341F-long and Bakt_805R-long in a dual indexed PCR approach (Klindworth, Pruesse, & Schweer, 2012). All PCR results, including negative controls, were purified using the AMPure bead calibration method, quantified using a fluorometric kit (QuantIT PicoGreen; Invitrogen), pooled in equimolar amounts, and sequenced paired-end using Illumina MiSeq at the Plate-forme d'analyses génomiques (IBIS, Université Laval).

2.4 | Amplicon analysis

Raw forward and reverse reads were quality trimmed, assembled into contigs for each sample, and classified using Mothur v.1.36.0 following the protocol of MiSeq SOP (https://www.mothur.org/wiki/MiSeq_SOP) (Kozich, Westcott, Baxter, Highlander, & Schloss, 2013; Schloss et al., 2009). Contigs were quality trimmed using several criteria. First, a maximum of two mismatches were allowed when aligning paired ends and ambiguous bases were excluded. Second, homo-polymers of more than eight, sequences with lengths <400 bp and >450 bp, sequences from chloroplasts, mitochondria, and nonbacterial were removed. Thirdly, chimeric sequences were found and removed using the UCHIME algorithm (Edgar, Haas, Clemente, Quince, & Knight, 2011). Moreover, the database SILVA was used for the alignment and the database RDP (v9) was used to classify the sequences with a 0.03 cutoff level. The Good's coverage index, which was used to evaluate the quality of the sequencing depth, was estimated in Mothur (Hurlbert, 1971).

2.5 | Statistical analyses

The analyses of microbiota were performed with Mothur and Rstudio v3.3.1 (RStudio Team, 2015). We first constructed a matrix of taxonomic composition (wild and captive included) with the number of operational taxonomic units (OTUs) after merging them by genus. The bacterial genera were considered as variables and fish as objects according to Mothur taxonomy files.

Details of the statistical analyses to test the effect of captivity (wild and captive conditions), the intestinal microbiota variation within and among wild whitefish populations as well as among the captive whitefish groups are presented in Appendix S1. In brief, a Spearman correlation matrix following a Hellinger transformation on the matrix of taxonomic composition was performed to document interactions between all captive and wild whitefish microbiota. The PERMANOVA analysis (number of permutations = 10,000) was also performed using the vegan package (Oksanen, Kindt, Legendre, & O'hara B., Stevens H.H., 2006) in R (Rstudio Team, 2015) on the matrix of taxonomic composition following a Hellinger transformation. An ANOVA following a fitted Gaussian family generalized model (GLM) was also performed at the alpha diversity level (inverse Simpson diversity) (Magurran, 2004). Furthermore, principal coordinates analyses (PCoAs) were built on a Bray–Curtis distance matrix after a Hellinger transformation to visualize variation between dwarf and normal whitefish within and among lakes (Legendre & Legendre, 1998; Oksanen et al., 2006). Finally, we documented the bacterial core of whitefish by identifying the bacterial genera present in 80% of all fish.

A linear discriminant analysis (LDA) was performed on the wild whitefish data, validated both according to (Evin et al., 2013) and from the PCA axes explaining at least 1% of the variation. The principal component analysis (PCA) was performed on the transformed Hellinger matrix.

In order to test for the presence of bacterial genera that were private to any of the captive whitefish group, we used the Metastats software with standard parameters ($p \leq .05$ and number of permutations = 1,000) to detect differential abundance of bacteria at the genus level between two host populations (White, Nagarajan, & Pop, 2009). Four Metastats analyses were performed on the captive whitefish between: dwarf versus normal, dwarf versus hybrid F1 D♀N♂, normal versus hybrid F1 N♀D♂, and hybrid F1 D♀N♂ versus hybrid F1 N♀D♂.

3 | RESULTS

3.1 | Sequencing quality

A total of 2,498,271 sequences were obtained after trimming for the entire data set composed of 185 whitefish intestinal microbiota (67 dwarf whitefish, 79 normal whitefish, and 39 hybrids whitefish) from wild and captive populations (Table S1). A total of 189,683 OTUs were identified with a 97% identity threshold, representing 710 bacterial genera.

The average Good's coverage estimation for all intestinal microbiota (wild and captive whitefish) was 92.3 ± 7.6%. This apparently low Good's coverage essentially came from captive whitefish microbiota with a mixed diet of *Artemia* and dry food (n = 47), with a coverage index of 82.8 ± 3.4%. Indeed, the Good's coverage from wild whitefish microbiota (n = 111) and captive whitefish microbiota with a diet of *Artemia* only (n = 27) were, respectively, 95.4 ± 2.8% and 98.2 ± 1.4%, thus indicating a good sequencing quality of our data. These data were considered reliable for further analyses for three reasons. First, the mixed diet captive group was composed of 341 bacterial genera in which the distribution showed an unusual high abundance (i.e., number of reads) for a few genera (Table S2), which is known to decrease the Good's coverage which is defined as 1-(Number of OTUs that have been sampled once/total number of sequences) (Hurlbert, 1971). Second, the Illumina MiSeq sequencing was performed in the same run for all samples, thus supporting the absence of sequencing problem given the excellent coverage obtained for the other groups. Third, a low Good's coverage is supposed to reflect a low number of sequences per sample because of the different filtration steps which eliminated reads generated by poor quality sequencing. Here, the low Good's coverage observed in the captive group that fed on a mixed diet showed a total number of sequences per sample similar to the other captive group (Table S2).

3.2 | Wild versus captive whitefish intestinal microbiota

The network analysis among all samples revealed a pronounced differentiation in intestinal microbiota between wild and captive whitefish (Figure 2). More specifically, all wild whitefish was comprised in a first group except one dwarf and two normal all from East Lake. There was no clear pattern of differentiation between wild dwarf and normal whitefish microbiota (Figure S1) but all wild populations

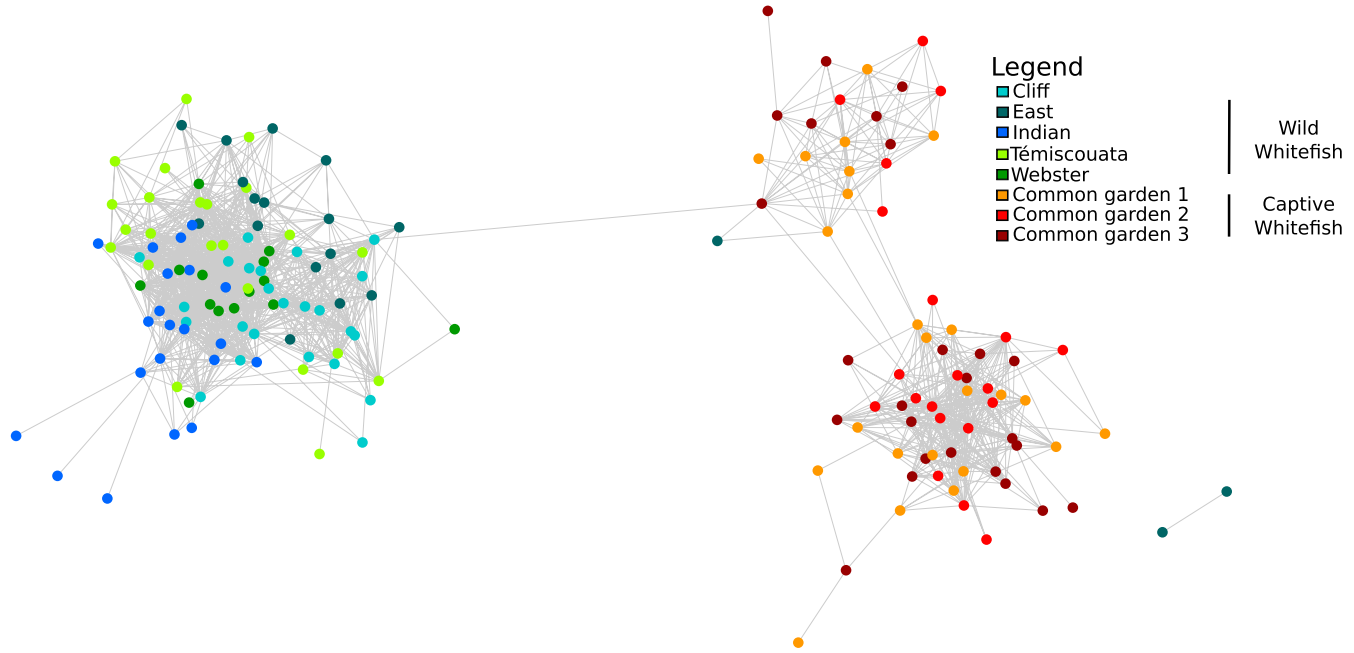


FIGURE 2 Network analysis of intestinal microbiota of dwarf and normal wild whitefish and intestinal microbiota of dwarf, normal, and hybrids captive whitefish. Each node represents either a dwarf, normal, or hybrid whitefish microbiota. The connecting lines between two samples represent their Spearman index correlation

tended to cluster distinctively from captive fish. The second and third groups were composed by all captive whitefish with few interactions observed between them despite the fact that they both comprised fish from all four groups (dwarf, normal and both reciprocal hybrids). This second level of differentiation was based on diet variation between the two captive groups (Figure 2). The differentiation between the wild and the captive fish was also supported by a significant effect of captivity on taxonomic composition (PERMANOVA, $p < .001$; Table 2) when performing analysis using all fish, dwarf only, and normal only, as well as on alpha diversity when using all fish (ANOVA, $p < .001$; Table S3). Furthermore, although the major phyla (*Firmicutes*, *Proteobacteria*, *Actinobacteria*, and *Planctomycetes*) were similar between wild and captive whitefish, the bacterial abundance clearly differed between them (Figure 3). Finally, among the 710 bacterial genera found among all captive and wild whitefish, six were shared by all fish: *Acinetobacter*, *Aeromonas*, *Clostridium*, *Legionella*, *Methylobacterium*, and *Propionibacterium*. These constitute the core intestinal microbiota defined as the microbial component shared by 80% of the samples.

3.3 | Wild dwarf and normal whitefish microbiota

At the phylum level, dwarf and normal wild whitefish transient intestinal microbiota was characterized by identical dominant phyla with a similar bacterial abundance (Figure 3). However, variation in taxonomic composition between dwarf and normal whitefish was observed for less dominant phyla. For example, *Tenericutes* and *Fusobacteria* were more represented in normal, whereas *Bacteroidetes* was more represented in dwarf whitefish. We observed a more pronounced influence of the lake of origin on taxonomic composition whereby dwarf or normal microbiota within a given lake shared more similarities than microbiota from different lake populations within a same species ($PERMANOVA_{lake}$, $p < .001$; $PERMANOVA_{species}$ $p < .006$; Table 2) (Figure 3).

Although no effect of lake or species on alpha diversity was observed (Table S3), there was a significant effect of both lake and host species on taxonomic composition (Table 2). The LDA performed on all wild whitefish also confirmed this overall difference between dwarf and normal intestinal microbiota albeit with overlap between them (Figure 4). Within each lake, the PERMANOVA revealed significant differences between dwarf and normal whitefish in three lakes (Cliff, East, and Indian lakes) but no difference in Témiscouata and Webster lakes (Table 2). Again, this suggested that the lake effect was more important than that of the host species. This was also supported by the PCoA analyses that revealed no global differentiation between all dwarf and normal whitefish (Figure 5a). Yet, host effect was supported in lake-specific PCoAs based on partially overlapping 95% confidence interval in Cliff, East, and Indian lakes (Figure 5b,d). Complete overlap was observed in Témiscouata Lake (Figure 5e), whereas results were ambiguous in Webster Lake, most likely due to low sample size for this lake (Figure 5f).

3.4 | Pure and hybrid whitefish microbiota in controlled environment

Although all fish were exposed to the same environment and the same food (both *Artemia* and dry fish food), we observed that some whitefish did not feed on the dry fish food and ate only live *Artemia*. As a result, we observed a mass and body length dichotomy between the two diet groups (Test of student, $p < .001$)

TABLE 2 Summary of PERMANOVA test statistics on microbiota taxonomic composition

Fish group	Source of variation	PERMANOVA		
		F-value	R^2	*p*(>*F*)
Wild				
All lakes	Species	2.350	.017	.006
	Lake	6.744	.197	<.001
	Species:Lake	1.927	.056	<.001
	Body mass	1.628	.012	.067
Cliff Lake	Species	5.253	.180	<.001
	Body mass	2.914	.100	<.001
East Lake	Species	1.889	.085	.047
	Body mass	1.165	.053	.291
Indian Lake	Species	2.032	.083	.041
	Body mass	1.582	.064	.105
Témiscouata Lake	Species	0.741	.033	.732
	Body mass	0.920	.041	.447
Webster Lake	Species	0.858	.057	.562
	Body mass	2.142	.143	.015
Captive	Group	1.985	.043	
	Diet	58.955	.427	<.001
	Species:Diet	1.557	.034	.108
	Body mass	1.990	.014	.084
	Tank	1.649	.024	.102
Both				
All fish groups	Captivity	64.457	.260	<.001
	Body mass	3.481	.014	.001
Dwarf	Captivity	28.245	.289	<.001
	Body mass	4.517	.046	<.001
Normal	Captivity	16.371	.180	<.001
	Body mass	1.917	.021	.035

Note: First, the fish group "wild" refers to the analysis of effect of host species (dwarf and normal), lake (Cliff, East, Indian, Témiscouata, and Webster), and its interaction with body mass as a covariate on all wild fish. Second, the fish group "all lakes" tests the host species and body mass as a covariate is treated for each lake separately. Third, the fish group "captive" refers to the analysis of effect of host group (dwarf, normal, hybrids F1 D♀N♂, and F1 N♀D♂), diet (*Artemia* only and mixed diet of live *Artemia* with dry food), and its interaction with body mass and tank as covariates on all captive fish. Fourth, the fish group "both" refers to the effect of captivity (wild and captive) and body mass as covariate on all fish, dwarf only, and normal only. *F*-value: value of the *F* statistic, R^2: *R*-squared statistic, *p*(>*F*): *p*-value. Only the interactions "Species:Lake" and "Species:Diet" are presented in this table.

(Table S4). As for the network analysis, the two distinct diet groups were evidenced by a significant effect of diet on both taxonomic composition microbiotas (PERMANOVA, $p < .001$; Table 2) and alpha diversity (ANOVA, $p = .001$; Table S3). The PCoA analysis clearly separated two distinct clusters on axis one corresponding to the two diet groups and independent of the genetic background (either pure forms or hybrids) (Figure 6). Furthermore, the mixed diet group was dominated by *Firmicutes* and the *Artemia* diet group was dominated by *Proteobacteria* (Figure 3). Within the mixed diet group, lower abundance for *Firmicutes*, but higher for *Proteobacteria*, was observed in reciprocal hybrids in comparison with dwarf and normal whitefish, whereas the opposite pattern was observed for the *Artemia* diet group (i.e., hybrids bacterial abundance was higher for *Firmicutes* but lower for *Proteobacteria*). Host group effect was also supported by the PERMANOVA (Table 2). The PCoA analysis within each of the two diet groups highlighted a modest differentiation based on overlapping 95% confidence interval between hybrids and pure whitefish (Figure 6). In the mixed diet group, dwarf and normal ellipses were mostly aligned on the second axis, whereas the ellipses of the two hybrid groups were mostly aligned on the first axis. The inverse pattern was observed in the *Artemia* diet group with the ellipses of the pure whitefish those of hybrid whitefish aligned on the first and second axes, respectively.

Between eight and 42 bacterial genera were differentially represented to a given whitefish group within diet groups (Figure 7). We observed 21 dwarf-specific and 27 normal-specific bacterial genera, respectively, whereas the comparison between hybrids F1 D♀N♂ and F1 N♀D♂ revealed 41 and 16 specific bacterial genera, respectively. Finally, we observed 135 specific bacteria genera in the mixed diet group versus 62 in the *Artemia* diet group (see Table S5 for details).

4 | DISCUSSION

4.1 | The intestinal microbiota of captive versus wild whitefish

Although an important part of bacteria which colonizes fish intestine may represent a random sampling from water and food, the occurrence of intestinal microbiota cores has been increasingly documented (Astudillo-García et al., 2017). The intestinal microbiota cores represent OTUs or genera shared among closed host relatives. Thus, despite the fact that wild and captive whitefish studied here never shared a common environment (they grew in totally different waters), the comparison of their microbiota highlighted six genera shared by at least 80% of all samples. Interestingly, our intestinal core microbiota data represented 20% of shared sequences which is higher than the intestinal microbiota core reported for cichlid species (13%–15%) (Baldo et al., 2015). These shared genera could be horizontally transmitted and/or selected as a common set of bacteria (Baldo et al., 2015; Rawls, Mahowald, Ley, & Gordon, 2006). Although the captive whitefish were hatched in captivity, their parents were of wild origin. Therefore, the conservation of certain genera by many captive whitefish might corroborate the microbiota vertical transmission in fish. It is also noteworthy that we found many bacteria of unknown taxonomy (see Figure 3) and much more so in wild than in captive whitefish. This, along with previous studies emphasizes the fact that a considerable number of bacteria are waiting to be discovered in natural freshwater ecosystems.

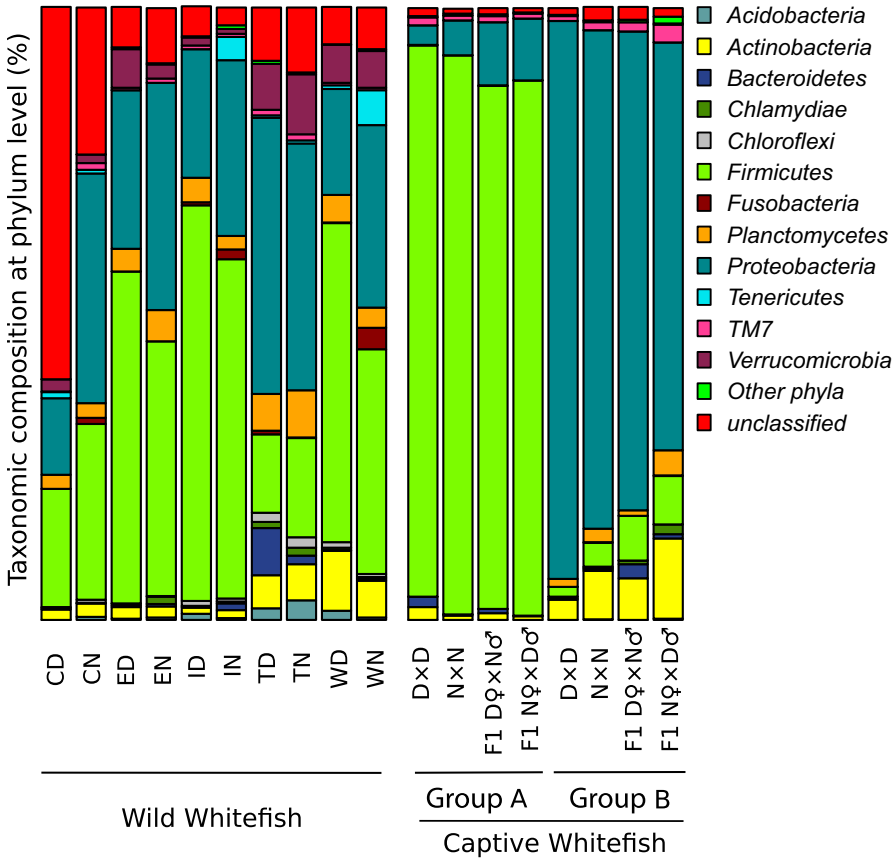


FIGURE 3 Relative abundance of phyla representatives found in intestinal microbiota for dwarf and normal wild whitefish in each lake, as well as in intestinal microbiota for dwarf, normal, and hybrids whitefish in controlled condition. Taxonomy was constructed with the database Silva and MOTHUR with confidence threshold at 97%. For the wild whitefish, lakes are represented as C: Cliff, E: East, I: Indian, T: Témiscouata, W: Webster, and the whitefish species is represented as D: dwarf and N: normal. For the captive fish, normal whitefish, dwarf whitefish, and hybrids are represented as N × N, D × D, F1 D♀×N♂ and F1 N♀×D♂, respectively. Diet group A (*Artemia* + dry food) and B (*Artemia*)

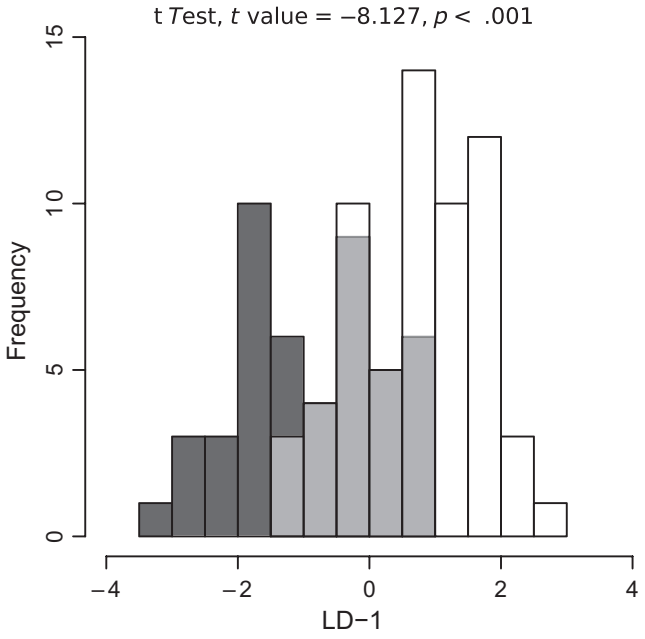


FIGURE 4 Linear discriminant analysis (LDA) histogram of all wild whitefish microbiota. This linear discriminant analysis was performed on the axes of principal component analysis (PCA) and *t* tests were performed on the results of the discriminant analysis. Dwarf and normal whitefish are represented by the black and white bars, respectively. Dwarf and normal whitefish with overlapping discriminant scores are shown in gray

4.2 | No clear pattern of parallel evolution in transient intestinal microbiota between dwarf and normal whitefish in the wild

Parallelism refers to the evolution of similar phenotypic traits in independent populations (Schluter & Nagel, 1995) and has been well documented in several sympatric species throughout the north hemisphere, including in Lake Whitefish (Bernatchez et al., 2010; Østbye et al., 2006; Schluter, 2000). Given the difference in trophic and ecologic niches occupied by both species (Landry & Bernatchez, 2010; Landry, Vincent, & Bernatchez, 2007), we predicted that some level of parallelism in transient intestinal microbiota would be observed between dwarf and normal whitefish species pairs. The dwarf whitefish is a limnetic fish feeding on zooplankton, whereas the normal whitefish is a benthic fish feeding on zoobenthos and molluscs (Bernatchez, Chouinard, & Lu, 1999; Bodaly, 1979). Therefore, we expected that a different diet should bring the dwarf and normal whitefish of a given sympatric pair in contact with different bacterial communities, leading to a distinct transient intestinal microbiota in a similar manner in the different lakes. Indeed, differentiation of microbiota composition correlated with diet was previously observed (David, Veena, & Kumaresan, 2016; Haygood & Jha, 2016; Koo et al., 2017; Nayak, 2010; Zarkasi et al., 2016). Thus, the use of novel diet elements may produce a change in the microbiota composition by increasing or decreasing different bacterial strain according to their metabolic potential (Rosenberg & Zilber, 2013). This is also supported by the microbiota composition differentiation of the two diet groups observed in captivity in this

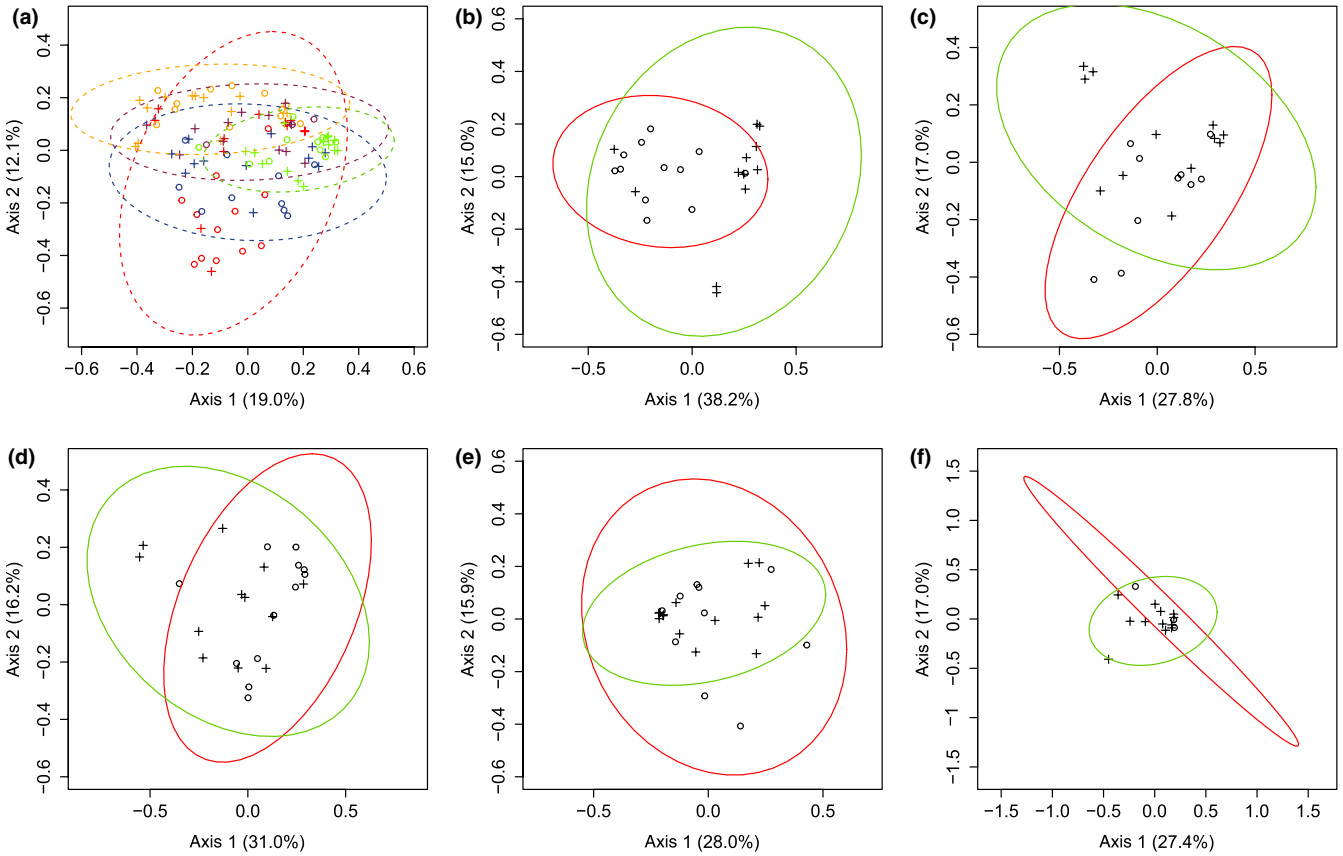


FIGURE 5 Principal coordinate analyses (PCoAs) within and between lakes for the wild whitefish microbiota. These PCoAs are based on Jaccard index after a Hellinger transformation. Ellipses of 95% confidence are illustrated and were done with dataEllips using R car package. (a) comparison among all wild whitefish populations from the five lakes. Each lake analyzed is represented by a different symbol and ellipse color: Cliff Lake (red), East Lake (blue), Indian Lake (orange), Témiscouata Lake (green), and Webster Lake (purple), and whitefish species is represented by symbols: Dwarf (circle) and Normal (cross). (b–f) comparison between Dwarf and Normal whitefish microbiota within each lake. Cliff Lake, East Lake, Indian Lake, Témiscouata Lake, and Webster Lake are represented by b, c, d, e, and f, respectively. Whitefish species is represented by different symbols: dwarf (circle) and normal (cross); ellipses of 95% confidence are illustrated and were done with dataEllips using R car package. The red and green ellipses represent the dwarf and normal species, respectively

study. Despite a global effect of species host on microbiota, we did not observe a clear pattern of parallelism among the five lakes comprising sympatric whitefish pairs studied here. Indeed, nonparallel difference between dwarf and normal whitefish microbiota composition was observed in three of the five lakes, whereas no difference was observed in the other two lakes. This indicated that the environment has a more pronounced effect than the species host on the transient intestinal microbiota of dwarf and normal whitefish. These results are in line with those obtained in a previous study in the same system but investigating kidney microbiota. Thus, Sevellec et al. (2014) showed that differences in bacteria composition between dwarf and normal whitefish were not parallel among lakes. However, unlike this study and in accordance with the higher diversity of prey types, normal whitefish kidney tissue consistently had a more diverse bacterial community and this pattern was parallel among lakes. Together, these results on whitefish microbiota add to building evidence from previous studies on this system that the adaptive divergence of dwarf and normal whitefish has been driven by both parallel and nonparallel ecological conditions across lakes, a situation reported in several other fishes (Oke, Rolshausen, LeBlond, & Hendry, 2017). Moreover, the water bacterial community of the same studied lakes was investigated previously and we found that each lake is characterized by a specific water bacterial community (Sevellec et al., 2018). This may reflect the differences in both biotic and abiotic factors among these lakes (Landry & Bernatchez, 2010; Landry et al., 2007). For instance, Cliff, Webster, and Indian lakes are characterized by a greater oxygen depletion and a lower zooplankton biomass, whereas East and Témiscouata lakes are characterized by more favorable environmental conditions with a more important biomass and broader size distribution of zooplanktonic prey and well-oxygenated water (Landry et al., 2007). Therefore, the variation in water bacterial community along with the biotic and abiotic factors could underlie the more important lake effect than species host effect observed in the transient intestinal microbiota. Nevertheless, highly distinct bacterial composition between the water bacterial community and the whitefish transient intestinal microbiota was observed among lakes. The water bacterial community was dominated by *Proteobacteria*, *Actinobacteria*, and *Bacteroidetes*, whereas the whitefish transient intestinal microbiota was dominated by *Firmicutes* and *Proteobacteria* (Sevellec et al., 2018). Therefore, whitefish transient intestinal microbiota was not directly

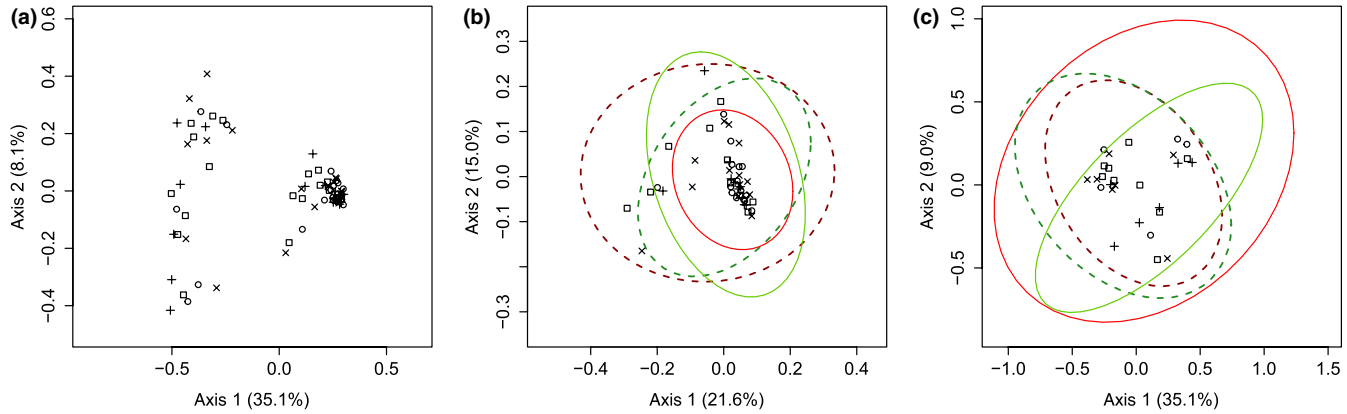


FIGURE 6 Principal coordinate analyses (PCoAs) between the microbiota of the four captive whitefish groups. (a) comparison between the four captive whitefish groups intestinal microbiota. (b) Comparison between the four whitefish groups intestinal microbiota in the mixed diet group. (c) Comparison between the four whitefish groups intestinal microbiota in the *Artemia* diet group. Ellipses of 95% confidence were done with dataEllips using R car package. Each whitefish species is represented by different symbols: dwarf (D♀×D♂), and normal (N♀×N♂) are represented by circle a cross respectively, and their ellipses are represented by continuous lines. The hybrid F1 N♀×D♂ and hybrid F1 D♀×N♂ are represented by the symbol × and □, respectively, and their ellipses are represented by dotted line. Dwarf and hybrid F1 D♀×N♂ are represented in red, whereas normal and hybrid F1 N♀×D♂ are represented in green

reflective of its local environment, which raises the hypothesis of a selective effect on microbiota induced by host physiology, immunity, and genetic background (Alberdi et al., 2016; Macke et al., 2017). For instance, some transient bacteria might contribute to digestion of host diet (Smith et al., 2015) and, in turn, may impact on the transient intestinal microbiota composition by increasing their abundance (Rosenberg & Zilber, 2013).

4.3 | Comparison of transient and adherent intestinal microbiota in wild whitefish and the host effect

The most prevalent phyla in wild whitefish transient microbiota are *Acidobacteria*, *Actinobacteria*, *Bacteroidetes*, *Chlamydiae*, *Chloroflexi*, *Firmicutes*, *Fusobacteria*, *Planctomycetes*, *Proteobacteria*, *Terenicutes*, *TM7*, and *Verrucomicrobia*, which have also been reported in previous studies of freshwater fishes (Eichmiller, Hamilton, Staley, Sadowsky, & Sorensen, 2016; Larsen & Mohammed, 2014; Li, Zhu, Yan, Ringø, & Yang, 2014; Roeselers et al., 2011; Sullam et al., 2012; Ye, Amberg, Chapman, Gaikowski, & Liu, 2014). In a previous study on adherent intestinal microbiota (that is adherent to the intestinal mucosa) performed on the same individuals, we found that while adherent and transient intestinal were characterized by similar major phyla, the abundance of some of them was different (Sevellec et al., 2018). For example, the five first phyla for the adherent microbiota were *Proteobacteria* (39.8%), *Firmicutes* (19%), *Actinobacteria* (5.1%), *OD1* (3.8%), and *Bacteroidetes* (2.8%), whereas the first five phyla for the transient microbiota were *Firmicutes* (38.2%), *Proteobacteria* (29.5%), *Verrucomicrobia* (4.4%), *Planctomycetes* (4.1%), and *Actinobacteria* (3.7%). Moreover, the number of genera and the number of OTUs were about 50% more important in the transient microbiota (611 genera and 94,883 OTUs) than the adherent microbiota (421 genera and 10,324 OTUs). Most of the adherent bacterial taxa living on the intestinal mucosa are not randomly acquired from the environment (Bolnick et al., 2014), but are rather retained by different host characteristics (Brucker & Bordenstein, 2012). Similarly, we previously reported that there is an important host effect in both dwarf and normal whitefish, which stabilizes the number of bacterial genera living in the intestinal mucosa (Sevellec et al., 2018). Thus, the comparison between whitefish transient and adherent microbiota supports the view that the whitefish host have a selective effect on its intestinal microbiota. For instance, dwarf and normal whitefish in Cliff and East lakes show a distinct intestinal microbiota for both the adherent and the transient bacteria, whereas the adherent, but not the transient intestinal microbiota differed between species in Témiscouata Lake, and the opposite was observed in Indian Lake. In Témiscouata Lake, this difference in adherent microbiota between species suggested a host species effect leading to differential abundance of the same bacterial taxa. In contrast, results in Indian Lake suggest that host species have no clear effect on microbiota divergence and that the difference in transient microbiota is likely caused by the trophic niches occupied by each species. Altogether, these observations suggest that the direction and intensity of factors determining the composition of intestinal microbiota may differ between the host and the microbiota of a given holobiont system, as previously reported (Rosenberg & Zilber, 2016). Here, we tentatively propose that three putative distinct host–microbiota interactions may have evolved independently in postglacial time: (a) divergence of intestinal microbiota influenced by the host and the environment (Cliff and East lakes), (b) divergence of the intestinal microbiota mostly influenced by the host (Témiscouata Lake), and (c) divergence of intestinal microbiota mostly influenced by the environment (Indian Lake). While speculative at this point, these putative distinct host–microbiota interactions would deserve to be carefully evaluated in future host–microbiota studies in a speciation context. Finally, given the pronounced difference that may exist between transient and adherent microbiota, our results suggest that

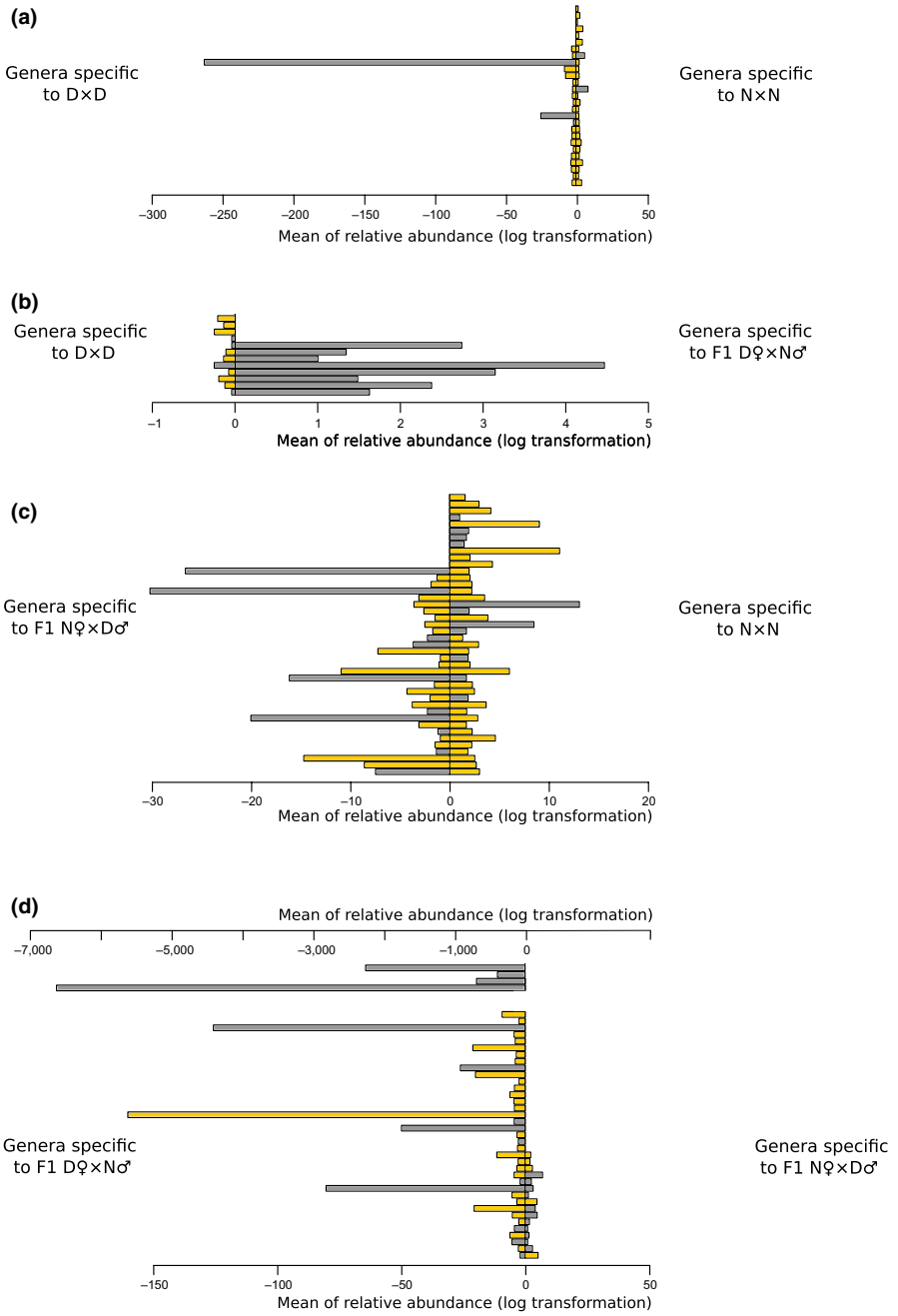


FIGURE 7 Metastats results for dwarf, normal, and hybrid captive whitefish. Four side-by-side comparisons were performed with dwarf (D♀×D♂), normal (N♀×N♂), hybrid F1 N♀×D♂, and hybrid F1 D♀×N♂. Each genus specific to a given whitefish group is represented by a bar plot. The abscissa represented the mean of the relative abundance of a genus specific after a log transformation. Mixed diet and *Artemia* diet groups are represented by yellow and gray bars, respectively

adherent microbiota is a more reliable choice to study the effect of host species on microbiota than the analysis of transient microbiota.

4.4 | Modest but significant host effect on the transient intestinal microbiota in controlled conditions

An unplanned variation in our experimental set up occurred during the captive rearing of the whitefish pair species and the reciprocal hybrids for seven months, which led to the unexpected observation of a diet preference which split the captive whitefish into two groups independently of the parental or hybrid origin or the tanks where fish were. The use of two types of food, *Artemia* and dry pellets, is usually recommended for optimizing growth and survival of juvenile whitefish in captivity (Flüchter, 1982; Zitzow & Millard, 1988). However, while 47 whitefish opted to feed on both types of diet, 27 chose to feed only on *Artemia*. Indeed, *Artemia* as the only source of food cannot provide the good nutrients used for the juvenile whitefish growth (Zitzow & Millard, 1988). As a result, all normal length whitefish were in the group A (dry food and *Artemia*), whereas all the small length whitefish were in the group B (*Artemia* only). (Table S4). We believe that it is very unlikely that any factor other than different diet would have caused such a strong association between size and microbiota. Indeed, this allowed us to assess the impact of different diets in an otherwise identical controlled environment, which revealed that diet had the most profound impact on the community composition of transient intestinal microbiota in a controlled environment.

Nevertheless, we did observe a significant, albeit modest effect of host groups on the transient intestinal microbiota. In principle,

in a controlled environment, there should be no environmental effect on the microbiota composition, and consequently, variation in microbiota should only depend on the host effect which integrated the influence of the host physiology, immunity, and genetic background. Here, while the PCoA analysis only revealed a slight pattern of differentiation between both parental species and their reciprocal hybrids, the PERMANOVA test revealed a statistically significant effect of the host genetic background on the taxonomic composition of the transient microbiota. This was accompanied by a significant variation in bacterial abundance at the phylum level, especially within the mixed diet group feeding on both *Artemia* and dry pellets. Finally, numerous genera that were specific to one whitefish species or the hybrids were observed in both diet groups. These results suggest an effect of hybrid genetic background on the transient intestinal microbiota. This effect could hypothetically be explained Bateson, Dobzhansky, and Muller (BDM) genetic incompatibilities previously documented in whitefish (Dion-Cote, Renaut, Normandeau, & Bernatchez, 2014; Gagnaire et al., 2013; Renaut, Nolte, & Bernatchez, 2009). To our knowledge, only one study compared the intestinal microbiota among closely related fish populations in controlled conditions (Sullam et al., 2015) and none compared parental and hybrid progeny. Specifically, distinct intestinal microbiota between two ecotypes of the Trinidadian Guppy (*Poecilia reticulata*) suggested a pronounced effect of the genetic background (Sullam et al., 2015). However, these results should be interpreted cautiously since fish used for this experiment were adults that were born in the wild and kept in tanks for 10 weeks only. Consequently, the difference could reflect a carry-over effect from the natural conditions, whereas in our case, fish were born in captivity.

To conclude, our results show that the transient intestinal fish microbiota is the result of complex interactions between the host's genetic background and environmental conditions. The prevalent environmental effect on the microbiota we observed among five sympatric whitefish pairs in the wild illustrates that drawing generalization regarding host–microbiota association for a given species may be difficult, and in fact inappropriate.

ACKNOWLEDGMENTS

We thank G. Côté, A. Dalziel, A-M. Dion-Côté, S. Higgins, and J-C Therrien for fieldwork and technical assistance for crossing and rearing of captive whitefish at the LARSA. We are grateful to C. Hernandez-Chàvez for laboratory advice and support and B. Boyle for his help with the Illumina MiSeq sequencing. We also thank the associate editor and two referees for their constructive input on an earlier version of the manuscript. This research was supported by a Discovery research grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to L.B. L.B also holds the Canadian Research Chair in genomics and conservation of aquatic resources.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

MS and LB conceived the study. LB contributed resources. MS and ML collected samples and analyzed the data. AB and MS performed laboratory assays. ND provided bioinformatics support (network analyses). MS performed bioinformatic work and wrote the manuscript, while LB, ML, and ND helped to draft and improve the manuscript. All authors edited the manuscript and approved the final version.

DATA AVAILABILITY STATEMENT

Sequencing results are available in the Sequence Read Archive (SRA) database at NCBI under BioProject ID SUB3062520.

ORCID

Maelle Sevellec https://orcid.org/0000-0002-2111-6702

REFERENCES

Alberdi, A., Aizpurua, O., Bohmann, K., Zepeda-Mendoza, M. L., & Gilbert, T. P. (2016). Do vertebrate gut metagenomes confer rapid ecological adaptation? *Trends in Ecology & Evolution*, *31*, 689–699. https://doi.org/10.1016/j.tree.2016.06.008

Astudillo-García, C., Bell, J. J., Webster, N. S., Glasl, B., Jompa, J., Montoya, J. M., & Taylor, M. W. (2017). Evaluating the core microbiota in complex communities: A systematic investigation. *Environmental Microbiology*, *19*, 1450. https://doi.org/10.1111/1462-2920.13647

Baldo, L., Pretus, J. L., Riera, J. L., Musilova, Z., Nyom, A. E. B., & Salzburger, W. (2017). Convergence of gut microbiotas in the adaptive radiations of African cichlid fishes. *The ISME Journal*, *11*, 1975–1987. https://doi.org/10.1038/ismej.2017.62

Baldo, L., Riera, J. L., Tooming-Klunderud, A., Albà, M. M., & Salzburger, W. (2015). Gut microbiota dynamics during dietary shift in eastern African Cichlid Fishes. *PLoS ONE*, *10*, e0127462. https://doi.org/10.1371/journal.pone.0127462

Béné, C., Barange, M., Subasinghe, R., Pinstrup-Andersen, P., Merino, G., Hemre, G. I., & Williams, M. (2015). Feeding 9 billion by 2050–Putting fish back on the menu. *Food Security*, *7*, 261–274. https://doi.org/10.1007/s12571-015-0427-z

Bernatchez, L., Chouinard, A., & Lu, G. (1999). Integrating molecular genetics and ecology in studies of adaptive radiation: Whitefish, *Coregonus* sp., as a case study. *Biological Journal of the Linnean Society*, *68*, 173–194. https://doi.org/10.1111/j.1095-8312.1999.tb01165.x

Bernatchez, L., Renaut, S., Whiteley, A. R., Derome, N., Jeukens, J., Landry, L., … St-Cyr, J. (2010). On the origin of species: Insights from the ecological genomics of Lake Whitefish. Philosophical transactions of the Royal Society of London. *Series B, Biological Sciences*, *365*, 1783–1800. https://doi.org/10.1098/rstb.2009.0274

Bodaly, R. A. (1979). Morphological and ecological divergence within the Lake Whitefish (*Coregonus clupeaformis*) species complex in Yukon Territory. *Journal of Fisheries Research Board of Canada*, *36*, 1214.

Bolnick, D. I., Snowberg, L. K., Hirsch, P. E., Christian, L., Lauber, C. L., Knight, R., … Svanbäck, R. (2014). Individuals' diet diversity influences gut microbial diversity in two freshwater fish (Threespine Stickleback and Eurasian Perch). *Ecology Letters*, *17*, 979–987. https://doi.org/10.1111/ele.12301

Brucker, R. M., & Bordenstein, S. R. (2012). Speciation by symbiosis. *Trends in Ecology & Evolution*, *27*, 443–451. https://doi.org/10.1016/j.tree.2012.03.011

Clements, K. D., Angert, E. R., Montgomery, W. L., & Choat, J. H. (2014). Intestinal microbiota in fishes: What's known and what's not. *Molecular Ecology*, *23*, 1891–1898. https://doi.org/10.1111/mec.12699

Dalziel, A. C., Laporte, M., Guderley, H., & Bernatchez, L. (2017). Do differences in the activities of carbohydrate metabolism enzymes between Lake Whitefish ecotypes match predictions from transcriptomic studies? *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, *224*, 138–149. https://doi.org/10.1016/j.cbpb.2017.08.001

Dalziel, A. C., Laporte, M., Rougeux, C., Guderley, H., & Bernatchez, L. (2016). Convergence in organ size but not energy metabolism enzyme activities among wild Lake Whitefish (*Coregonus clupeaformis*) species pairs. *Molecular Ecology*, *26*, 225–244.

Dalziel, A. C., Martin, N., Laporte, M., Guderley, H., & Bernatchez, L. (2015). Adaptation and acclimation of aerobic exercise physiology in Lake Whitefish ecotypes (*Coregonus clupeaformis*). *Evolution*, *69*, 2167–2186.

Damodaram, K., Ayyasamy, A., & Kempraj, V. (2016). Commensal bacteria aid mate-selection in the Fruit Fly, *Bactrocera dorsalis*. *Microbial Ecology*, *72*, 725–729. https://doi.org/10.1007/s00248-016-0819-4

David, C. J., Veena, R., & Kumaresan, G. (2016). High cholesterol diet induces obesity in Zebrafish. *PLoS ONE*, *8*, e66970.

David, L. A., Maurice, C. F., Carmody, R. N., Gootenberg, D. B., Button, J. E., Wolfe, B. E., … Turnbaugh, P. J. (2014). Diet rapidly and reproducibly alters the Human gut microbiome. *Nature*, *505*, 559–563. https://doi.org/10.1038/nature12820

Dion-Cote, A. M., Renaut, S., Normandeau, E., & Bernatchez, L. (2014). Article RNA-seq reveals transcriptomic shock involving transposable elements reactivation in hybrids of young Lake Whitefish species. *Molecular Biology and Evolution*, *31*, 1188.

Dobzhansky, T. (1937). *Genetics and the origin of species*. New York, NY: Columbia University Press.

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., & Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, *27*, 2194–2200. https://doi.org/10.1093/bioinformatics/btr381

Eichmiller, J. J., Hamilton, M. J., Staley, C., Sadowsky, M. J., & Sorensen, P. W. (2016). Environment shapes the fecal microbiome of invasive carp species. *Microbiome*, *4*, 44. https://doi.org/10.1186/s40168-016-0190-1

Evin, A., Cucchi, T., Cardini, A., Vidarsdottir, U. S., Greger, L. G., & Dobney, K. (2013). The long and winding road: Identifying pig domestication through molar size and shape. *Journal of Archaeological Science*, *40*, 735–743. https://doi.org/10.1016/j.jas.2012.08.005

Flüchter, J. (1982). Substance essential for metamorphosis of fish larvae extracted from *Artemia*. *Aquaculture*, *27*, 83. https://doi.org/10.1016/0044-8486(82)90113-2

Gagnaire, P. A., Normandeau, E., Pavey, S. A., & Bernatchez, L. (2013). Mapping phenotypic, expression and transmission ratio distortion QTL using RAD markers in the Lake Whitefish (*Coregonus clupeaformis*). *Molecular Ecology*, *22*, 3036–3048.

Guivier, E., Martin, J. F., Pech, N., Ungaro, A., Chappaz, R., & Gilles, A. (2017). Microbiota diversity within and between the tissues of two wild interbreeding species. *Microbial Ecology*, *75*, 799–810. https://doi.org/10.1007/s00248-017-1077-9

Hanif, A., Bakopoulos, V., & Dimitriadis, G. J. (2004). Maternal transfer of humoral specific and non-specific immune parameters to sea bream (*Sparus aurata*) larvae. *Fish & Shellfish Immunology*, *17*, 411–435. https://doi.org/10.1016/j.fsi.2004.04.013

Hata, H., Tanabe, A. S., Yamamoto, S., Toju, H., Kohda, M., & Hori, M. (2014). Diet disparity among sympatric herbivorous cichlids in the same ecomorphs in Lake Tanganyika: Amplicon pyrosequences on algal farms and stomach contents. *BMC Biology*, *12*, 90. https://doi.org/10.1186/s12915-014-0090-4

Haygood, A. M., & Jha, R. (2016). Strategies to modulate the intestinal microbiota of Tilapia (*Oreochromis* sp.) in aquaculture: A review. *Reviews in Aquaculture*, *1*, 1–14.

Hooper, L. V., Wong, M. H., Thelin, A., Hansson, L., Falk, P. G., & Gordon, J. I. (2001). Molecular analysis of commensal host-microbial relationships in the intestine. *Science*, *291*, 881–884. https://doi.org/10.1126/science.291.5505.881

Hurlbert, S. H. (1971). The nonconcept of species diversity: A critique and alternative parameters. *Ecological Society of America*, *52*, 577–586. https://doi.org/10.2307/1934145

Jeukens, J., Bittner, D., Knudsen, R., & Bernatchez, L. (2009). Candidate genes and adaptive radiation: Insights from transcriptional adaptation to the limnetic niche among coregonine fishes (*Coregonus* spp., *Salmonidae*). *Molecular Biology and Evolution*, *26*, 155–166. https://doi.org/10.1093/molbev/msn235

Klindworth, A., Pruesse, E., & Schweer, T. (2012). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Research*, *41*, e1.

Koo, H., Hakim, J. A., Powell, M. L., Kumar, R., Eipers, P. G., Morrow, C. D., … Bej, A. K. (2017). Metagenomics approach to the study of the gut microbiome structure and function in zebrafish *Danio rerio* fed with gluten formulated diet. *Journal of Microbiological Methods*, *135*, 69–76. https://doi.org/10.1016/j.mimet.2017.01.016

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology*, *79*, 5112–5120. https://doi.org/10.1128/AEM.01043-13

Landry, L., & Bernatchez, L. (2010). Role of epibenthic resource opportunities in the parallel evolution of Lake Whitefish species pairs (*Coregonus* sp.). *Journal of Evolutionary Biology*, *23*, 2602–2613. https://doi.org/10.1111/j.1420-9101.2010.02121.x

Landry, L., Vincent, W. F., & Bernatchez, L. (2007). Parallel evolution of Lake Whitefish dwarf ecotypes in association with limnological features of their adaptive landscape. *Journal of Evolutionary Biology*, *20*, 971–984. https://doi.org/10.1111/j.1420-9101.2007.01304.x

Laporte, M., Dalziel, A. C., Martin, N., & Bernatchez, L. (2016). Adaptation and acclimation of traits associated with swimming capacity in Lake Whitefish (*Coregonus clupeaformis*) ecotypes. *Journal of Evolutionary Biology*, *16*, 160. https://doi.org/10.1186/s12862-016-0732-y

Laporte, M., Rogers, S. M., Dion, A. M., Normandeau, E., Gagnaire, P. A., Dalziel, A. C., … Bernatchez, L. (2015). RAD-QTL mapping reveals both genome-level parallelism and different genetic architecture underlying the evolution of body shape in Lake Whitefish (*Coregonus clupeaformis*) Species Pairs. *G3: Genes | Genomes | Genetics*, *5*, 1481.

Larsen, A. M., Mohammed, H. H., & Arias, C. R. (2014). Characterization of the gut microbiota of three commercially valuable warmwater fish species. *Journal of Applied Microbiology*, *116*, 1396. https://doi.org/10.1111/jam.12475

Legendre, P., & Legendre, L. (1998). *Numerical ecology* (2nd ed.). Amsterdam, the Netherlands: Elsevier Science.

Li, X. M., Zhu, Y. J., Yan, Q. Y., Ringø, E., & Yang, D. G. (2014). Do the intestinal microbiotas differ between paddlefish (*Polyodon spathala*) and bighead carp (*Aristichthys nobilis*) reared in the same pond? *Journal of Applied Microbiology*, *117*, 1245–1252.

Llewellyn, M. S., Boutin, S., Hoseinifar, S. H., & Derome, N. (2014). Teleost microbiomes: The state of the art in their characterization, manipulation and importance in aquaculture and fisheries. *Frontiers in Microbiology*, *5*, 207. https://doi.org/10.3389/fmicb.2014.00207

Macke, E., Tasiemski, A., Massol, F., Callens, M., & Decaestecker, E. (2017). Life history and eco-evolutionary dynamics in light of the gut microbiota. *Oikos*, *126*, 508–531. https://doi.org/10.1111/oik.03900

Magurran, A. E. (2004). *Measuring biological diversity*. Oxford, UK: Blackwells.

Miller, E. T., Svanbäck, R., & Bohannan, B. (2018). Microbiomes as metacommunities: Understanding host-associated microbes through metacommunity ecology. *Trends in Ecology & Evolution*, *33*(12), 926–935. https://doi.org/10.1016/j.tree.2018.09.002

Muller, H. J. (1942). Isolating mechanisms, evolution and temperature. *Biology Symposium*, *6*, 71–125.

Nayak, S. K. (2010). Role of gastrointestinal microbiota in fish. *Aquaculture Research*, *41*, 1553–1573. https://doi.org/10.1111/j.1365-2109.2010.02546.x

Nelson, J. S. (2006). *Fishes of the world* (4th ed.). New York, NY: Wiley.

Oke, K. B., Rolshausen, G., LeBlond, C., & Hendry, A. P. (2017). How parallel is parallel evolution? A comparative analysis in fishes. *The American Naturalist*, *190*, 1–16. https://doi.org/10.1086/691989

Oksanen, J., Kindt, R., Legendre, P., O'hara, B., & Stevens, H. H. (2006). *Vegan: community ecology package ver. 2.0–10*.

Østbye, K., Amundsen, P. A., Bernatchez, L., Klemetsen, A., Knudsen, R., Kristoffersen, K., … Hindar, K. (2006). Parallel evolution of ecomorphological traits in the European whitefish *Coregonus lavaretus* (L.) species complex during post-glacial times. *Molecular Ecology*, *15*, 3983–4001.

Rawls, J. F., Mahowald, M. A., Ley, R. E., & Gordon, J. I. (2006). Reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients reveal host habitat selection. *Cell*, *127*, 423–433. https://doi.org/10.1016/j.cell.2006.08.043

Rawls, J. F., Samuel, B. S., & Gordon, J. I. (2004). Gnotobiotic zebrafish reveal evolutionarily conserved responses to the gut microbiota. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 4596–4601.

Renaut, S., Nolte, A. W., & Bernatchez, L. (2009). Gene expression divergence and hybrid misexpression between Lake Whitefish species pairs (*Coregonus* spp. *Salmonidae*). *Molecular Biology and Evolution*, *26*, 925–936. https://doi.org/10.1093/molbev/msp017

Roeselers, G., Mittge, E. K., Stephens, W. Z., Parichy, D. M., Cavanaugh, C. M., Guillemin, K., & Rawls, J. F. (2011). Evidence for a core gut microbiota in the zebrafish. *The ISME Journal*, *5*, 1595–1608. https://doi.org/10.1038/ismej.2011.38

Rosenberg, E., & Zilber, I. (2013). *The hologenome concept: Human, animal and plant microbiota*. Basel, Switzerland: Springer.

Rosenberg, E., & Zilber, I. (2016). Microbes drive evolution of animals and plants: The hologenome concept. *MBio*, *7*, e01395–e1415. https://doi.org/10.1128/mBio.01395-15

Rougeux, C., Bernatchez, L., & Gagnaire, P. A. (2017). Modeling the multiple facets of speciation-with-gene-flow toward inferring the divergence history of Lake Whitefish species pairs (*Coregonus clupeaformis*). *Genome Biology and Evolution*, *9*, 2057–2074. https://doi.org/10.1093/gbe/evx150

RStudio Team (2015). *RStudio: Integrated Development for R*. Boston, MA: RStudio Inc.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., … Weber, C. F. (2009). Introducing Mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, *75*, 7537–7541. https://doi.org/10.1128/AEM.01541-09

Schluter, D. (2000). Ecological character displacement in adaptive radiation. *The American Naturalist*, *156*, 4–16. https://doi.org/10.1086/303412

Schluter, D., & Nagel, L. M. (1995). Parallel speciation by natural selection. *The American Naturalist*, *146*, 292–301. https://doi.org/10.1086/285799

Sevellec, M., Derome, N., & Bernatchez, L. (2018). Holobionts and ecological speciation: The intestinal microbiota of Lake Whitefish species pairs. *Microbiome*, *6*, 47. https://doi.org/10.1186/s40168-018-0427-2

Sevellec, M., Pavey, S. A., Boutin, S., Filteau, M., Derome, N., & Bernatchez, L. (2014). Microbiome investigation in the ecological speciation context of Lake Whitefish (*Coregonus clupeaformis*) using next-generation sequencing. *Journal of Evolutionary Biology*, *27*, 1029.

Shropshire, J. D., & Bordenstein, S. R. (2016). Speciation by symbiosis: The microbiome and behavior. *MBio*, *7*, e01785–e1815. https://doi.org/10.1128/mBio.01785-15

Smith, C., Snowberg, L. K., Caporaso, J. G., Knight, R., & Bolnick, D. I. (2015). Dietary input of microbes and host genetic variation shape among-population differences in stickleback gut microbiota. *The ISME Journal*, *9*, 2515–2526. https://doi.org/10.1038/ismej.2015.64

StCyr, J., Derome, N., & Bernatchez, L. (2008). The transcriptomics of life-history trade-offs in whitefish species pairs (*Coregonus* sp.). *Molecular Ecology*, *17*, 850–870.

Sullam, K. E., Essinger, S. D., Lozupone, C. A., O'connor, M. P., Rosen, G. L., Knight, R., … Russell, J. A. (2012). Environmental and ecological factors that shape the gut bacterial communities of fish: A meta-analysis. *Molecular Ecology*, *21*, 3363–3378. https://doi.org/10.1111/j.1365-294X.2012.05552.x

Sullam, K. E., Rubin, B., Dalton, C. M., Kilham, S. S., Flecker, A. S., & Russell, J. A. (2015). Divergence across diet, time and populations rules out parallel evolution in the gut microbiomes of Trinidadian guppies. *The ISME Journal*, *9*, 1508–1522. https://doi.org/10.1038/ismej.2014.231

White, J. R., Nagarajan, N., & Pop, M. (2009). Statistical methods for detecting differentially abundant features in clinical metagenomic samples. *PLoS Computational Biology*, *5*, e1000352. https://doi.org/10.1371/journal.pcbi.1000352

Wilkins, L., Rogivue, A., Fumagalli, L., & Wedekind, C. (2015). Declining diversity of egg-associated bacteria during development of naturally spawned whitefish embryos (*Coregonus* spp.). *Aquatic Sciences*, *77*, 481–497. https://doi.org/10.1007/s00027-015-0392-9

Woese C. R. (1998). Default taxonomy: Ernst Mayr's view of the microbial world. *Proceedings of the National Academy of Sciences of the United States of America*, *95*, 11043–11046. https://doi.org/10.1073/pnas.95.19.11043

Ye, L., Amberg, J., Chapman, D., Gaikowski, M., & Liu, W. T. (2014). Fish gut microbiota analysis differentiates physiology and behavior of invasive Asian carp and indigenous American fish. *The ISME Journal*, *10*, 2076.

Zarkasi, K. Z., Taylor, R. S., Abell, G., Tamplin, M. L., Glencross, B. D., & Bowman, J. P. (2016). Atlantic salmon (*Salmo salar* L.) gastrointestinal microbial community dynamics in relation to digesta properties and diet. *Molecular Ecology*, *71*, 589.

Zitzow, R. E., & Millard, J. L. (1988). Survival and growth of Lake Whitefish (*Coregonus clupeaformis*) larvae fed only formulated dry diets. *Aquaculture*, *69*, 105. https://doi.org/10.1016/0044-8486(88)90190-1

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Sevellec M, Laporte M, Bernatchez A, Derome N, Bernatchez L. Evidence for host effect on the intestinal microbiota of whitefish (*Coregonus* sp.) species pairs and their hybrids. *Ecol Evol*. 2019;9:11762–11774. https://doi.org/10.1002/ece3.5676

Appendix S1

Experimental crosses of captive whitefish

Whitefish eggs used for this study were incubated at the Laboratoire de Recherche en Sciences Aquatiques (LARSA, Université Laval, Québec, Canada). The dwarf species came from Témiscouata Lake (47°40'04"N, 68°49'03"W) which is from the Acadian glacial lineage origin whereas the normal species came from Aylmer Lake (45°54"N, 71°20"W) corresponding to the Atlantic glacial lineage (Bernatchez & Dodson, 1991). Backcross F1-Hybrids were obtained by crossing a F1 hybrids laboratory strain and wild whitefish parents. More precisely, F1 hybrid (F1 D♀*N♂) were produced in crossing three wild dwarf females and two laboratory strain normal males (born in laboratory from Aylmer Lake population origin) by artificial fertilization. Same processes was used to produced F1 hybrid (F1 N♀*D♂) with crossing five laboratory strain normal females and twelve wild dwarf males (see figure 1 (Rogers et al., 2007)). The dwarf and normal whitefish crosses were also created by artificial fertilization with sperm and eggs were collected in the field and transported to the LARSA. No treatments, such as antibiotics or malachite green were delivered to the eggs.

Whitefish microbiota: DNA extraction, amplification and sequencing

Slight modifications of the QIAmp© Fast DNA stool mini kit (QIAGEN) were performed to extract bacterial DNA captive and wild fish intestinal bolus. To maximize DNA extraction of gram-positive bacteria, temperature and time were increased during the incubation steps and all products used were doubled (Proteinase K, Buffer AL and ethanol 100%). Thus, 1200 µl were transferred into the column (in two subsequent steps) and bacterial DNA was eluted from the column with 100 µl of ultrapure water (DEPC-treated Water Ambion®). DNA extractions were

quantified with a Nanodrop (Thermo Scientific) and stored at −20°C until use. Five blank extractions were also done as negative controls.

The PCR amplification comprised 50 µl PCR amplification mix containing 25 µl of NEBNext Q5 Hot Start Hifi PCR Master Mix, 1 µl (0.2 µm) of each specific primers (Bakt_341F-long and Bakt_805R-long), 15 µl of sterile nuclease-free water and 8µl of specify amount of DNA. The PCR program consisted of an initial denaturation step at 98°C for 30s, followed by 30 cycles, where one cycle consisted of 98°C for 10 s (denaturation), 56°C for 30 s (annealing) and 72°C for 45s (extension), and a final extension of 72°C for 5 min. Negative and positive controls were also performed using the same program.

Statistical analyses

We first investigated the microbiota difference between the captive and wild whitefish using a network analysis. A Spearman's correlation matrix following a Hellinger transformation on the matrix of taxonomic composition was performed to document interactions between all captive and wild whitefish microbiota. More precisely, a Spearman's correlation value (threshold ≥0.5), a P-value and Bonferroni correction was calculated for each sample. The network was visualized using Cytoscape v3.2.1 (Shannon *et al.*, 2003), where nodes were illustrated in two different versions: (i) according to their sampling sites (eight groups: five lakes and three tanks) and (ii) according to their genetic group (the two wild species pairs and the four captive groups: dwarf, normal, reciprocal hybrid F1 D♀N♂, and hybrid F1 N♀D♂). We also tested for the effect of captivity (wild and captive conditions) on whitefish microbiota taxonomic composition (PERMANOVAs; 10,000 permutations) and alpha diversity (inverse Simpson diversity) with an ANOVA following a fitted Gaussian family generalized model (GLM) (Magurran, 2004). This was performed on all fish, on dwarf whitefish only and on normal whitefish only.

Secondly, in order to document variation within and among wild whitefish populations, we tested for an effect of 'host species', 'lake' and their interaction, with 'body mass' as a covariate on the taxonomic composition, using a permutational analysis of variance (PERMANOVA; 10,000 permutations). This procedure was run for each of the five lakes independently after removing the explanatory variable 'lake' of the analysis. The 'host species', 'lake' effects and their interaction on the inverse Simpson diversity were also tested using an analysis of variance (ANOVA) following a fitted Gaussian family generalized model (GLM). Allometric effect on inverse Simpson diversity was first tested with a linear regression on body mass. As for the taxonomic composition, we ran this procedure for each lake independently. Furthermore, principal coordinates analyses (PCoAs) was built on a Bray-Curtis distance matrix after a Hellinger transformation to visualize variation at the genus level between dwarf and normal wild whitefish among and within the lakes (Oksanen *et al.*, 2006; Legendre & Legendre, 1998).

Finally, we tested for differences in taxonomic composition between the four captive groups by investigating the effect of 'host group' (Dwarf, Normal, hybrid F1 D♀N♂, and hybrid F1 N♀D♂), 'diet' and their interaction with 'body mass' and 'tank' as covariates (PERMANOVA; 10,000 permutations). The effect of diet was added in the analysis because fish bolus exhibited a clear distinction between two observed feeding habits during the controlled experiment (A: feeding on a mix of dry food and *Artemia*, B: feeding on *Artemia* only). For the alpha diversity, the effect of 'host group', 'diet' and their interaction on the inverse Simpson diversity were tested with a mixed effects linear random model using the 'nlme' package in R, with tank as a random effect and individual fish nested within tank (Pinheiro *et al.*, 2009). As for the analyses on wild whitefish, we first tested for an allometric relationship with body mass using a linear regression and used the residuals in all cases showing a significant relationship. Principal coordinates

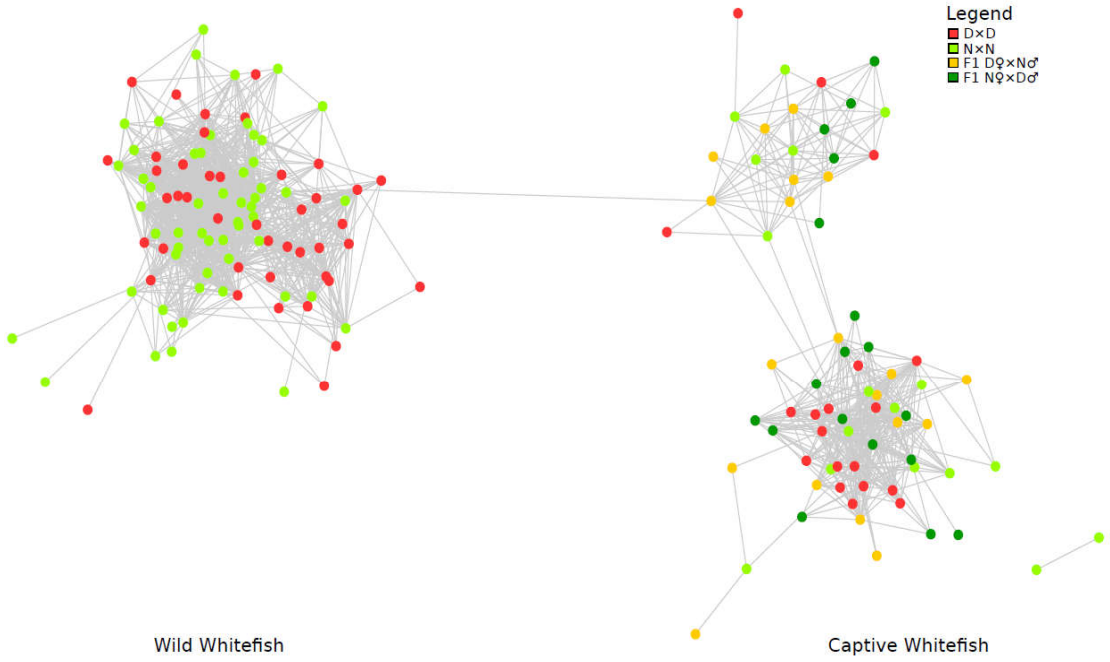
analyses (PCoAs) built on a Bray-Curtis distance matrix after a Hellinger transformation were also used to visualize variation at the genus level as described above. Linear discriminant analyses were also performed on captive whitefish but results were not displayed because of a negative cross-validation according to Evin *et al.*(2013).

Whitefish host: DNA extraction, amplification and genetic identification of captive whitefish lineages

A fin clip was collected from all fish and DNA was extracted using a salt extraction method (Aljanabi & Martinez, 1997) with slight modifications (Valiquette *et al.*, 2014). Mitochondrial (mtDNA) and nuclear DNA were used to identify the whitefish dwarf and normal, and their hybrids (F1 hybrid D♀N♂ and F1 hybrid N♀D♂). First, an analysis of mtDNA restriction fragment length polymorphism (RFLP) was performed as described in Dalziel et al. since pure dwarf and normal species possess distinct mitochondrial DNA haplotypes (Jacobsen *et al.*, 2012; Dalziel *et al.*, 2015). In brief, after the amplification of the cytochrome b by PCR, the amplified products were digested with SnaBI which cuts the amplified cytochrome b of the normal whitefish haplotype but not of the dwarf. Second, 12 nuclear microsatellite loci were genotyped on all juvenile whitefish and their known parents to differentiate them at the nuclear DNA level and details about primer sequences and PCR protocols are presented in Rico et al.. Three different PCRs were performed for this whitefish microsatellite markers analysis (Rico *et al.*, 2013). Firstly, the multiplex PCR A was performed with 2 µl (≈20 ng) of whitefish DNA, 5 µL Qiagen® multiplex reaction buffer, forward and reverse primers at different concentrations: 0.3 µm of Cocl32, Cocl lav41, Cocl Lav8 and 0.35 µm of Cocl Lav224; purified water adjusted the final volume at 10 µl. Multiplex PCR program was: 15 min at 94°C, and then 35 cycles of 30 sec at 94°C, 3 min at 58°C, 1 min at 72°C and 30 min at 60°C. Secondly, the multiplex PCR B were performed with 2 µl (≈20 ng) of whitefish DNA, 5 µL Qiagen® multiplex reaction buffer and forward and reverse primers at different concentration: 6 µm of Cocl15 et Cisco200 and 0.25

µm of Cocl 33; purified water adjusted the final volume at 10 µl. Multiplex PCR program was: 15 min at 94°C, and then 35 cycles of 30 sec at 94°C, 3 min at 60°C, 1 min at 72°C and 30 min at 60°C. Thirdly, the Simplex PCRs were performed with 2 µl (≈20 ng) whitefish DNA, 0.2 µl GoTaq® DNA polymerase (PROMEGA), 0.5 µl of each forward and reverse markers (0.5 µm) (Osmo5, Cocl34, Cocl36, Bwf F-1 and Cocl Lav22) 2 µl of 5X Colorless GoTaq®, 0.6 µl of MgCl2 (0.5 mM), 0.8 µl dNTPs (200 µm) and purified water adjusted the final volume at 10 µl. Simplex PCR program was: 2 min at 94°C, and then 35 cycles of 30 sec at 94°C, 3 min at 58°C (Osmo5, Cocl36, Bwf F-1, Cocl Lav22) or 64°C (Cocl34), 1 min at 72°C and 30 min at 60°C. Amplified loci were migrated via electrophoresis using an ABI 3130xl capillary DNA sequencer (Applied Biosystems Inc.) with a molecular size standard (GeneScan-500 LIZ, Applied Biosystems). Genotypes were scored using Genemapper 4.0 (Applied Biosystems Inc). A combination of three software, STRUCTURE v2.3.4, GENECLASS2 v2.0 and PAPA v2.0 was used to reassign each studied fish to its group of origin (Pritchard *et al.*, 2000; Duchesne *et al.*, 2002; Piry *et al.*, 2004). STRUCTURE was performed assuming an admixture model without priors with a burn-in period of 50 000 followed and 100,000 Markov Chain Monte Carlo (MCMC) steps. GENECLASS2 was conducted using the simulation test of (Rannala & Mountain, 1997) based on 100,000 simulated individuals. Finally, PAPA was performed for the parental allocation procedure with a uniform error model (error sum = 0.02).

Supplementary Figures



112

113 Figure S1 Network analysis of intestinal microbiota of dwarf and normal wild whitefish and
114 intestinal microbiota of dwarf, normal and hybrids captive whitefish. The nodes represent a dwarf
115 or a normal or a hybrid whitefish microbiota. More precisely, DD: dwarf whitefish, NN: normal
116 whitefish, DH: hybrid F1 D♀*N♂, NH: F1 N♀*D♂. The connecting lines between two samples
117 represent their correlation and is highlighting by a Spearman index.

118

Table S1 Steps used to reduce sequencing and PCR errors. We followed the step recommended by MOTHUR in the MiSeq SOP protocol.

Main Step of filtration	Number of microbiota wild whitefish sequences	Number of microbiota captive whitefish sequences	Number of microbiota captive and wild whitefish sequences
After contigs construction	4765128	2498271	8299965
Remove sequences with ambiguous bases and lengths more than 450 bp	1774729	885737	2660466
Aligning paired ends (maximum two mismatches) and remove sequences with homopolymers of more than eight bp and with lengths less than 400 bp	1737037	872921	2609958
Remove chimeric sequences	1729598	868868	2598519
Remove sequences from chloroplasts, mitochondria and nonbacterial	1855778	845370	2498271
Number of final sequences	1855778	845370	2498271
Number of OTUs	94883	85363	189683
Number of genus	611	433	710
Good's Coverage	95.06%	88.22%	91.86%

Table S2 Matrix of bacterial abundance and Good's coverage per captive whitefish sample. DD: dwarf whitefish, NN: normal whitefish, DH: hybrid F1 D♀*N♂, NH: F1 N♀*D♂. The diet group A is composed of Artemia and dry food; B is composed of Artemia.

Group	Whitefish species	Body mass (gram)	Diet Group	Number of sequences	Good's coverage
04-B2F	NH	62.50	A	9554	83.63
04-B3F	NH	68.45	A	4047	78.95
05-B1F	NN	54.11	A	7950	87.38
05-B2F	DD	48.75	A	4272	77.76
05-B3F	NH	49.62	A	10535	84.72
06-B1F	DH	20.27	A	10626	86.82
06-B2F	NH	53.55	A	9598	85.02
07-B1F	DD	70.76	A	10220	83.05
07-B3F	NN	66.45	A	12571	88.09
08-B3F	NH	38.24	A	4981	84.08
09-B1F	DD	35.43	A	8830	86.87
09-B3F	NN	78.10	A	9552	78.43
10-B1F	DD	66.88	A	8965	82.96
10-B3F	DD	40.19	A	5368	80.61
11-B1F	DD	70.33	A	9494	83.98
11-B3F	DD	15.23	A	4764	76.20
12-B1F	NH	41.27	A	2616	81.04
12-B2F	DH	12.24	A	11443	81.56
12-B3F	NH	96.84	A	8438	84.55
13-B2F	DH	15.65	A	8778	84.43
14-B1F	NH	45.23	A	7496	80.31
14-B3F	DD	61.29	A	8842	80.32
15-B1F	DH	26.03	A	9462	86.81
15-B2F	NN	97.68	A	9109	82.89
15-B3F	DH	51.04	A	7660	79.13
16-B3F	NH	70.71	A	7866	78.85
17-B1F	NN	64.96	A	4621	78.84
17-B3F	DD	86.65	A	7478	80.82
18-B2F	DD	78.20	A	7387	81.18
19-B3F	NN	83.91	A	14935	82.68
20-B1F	DD	22.63	A	9012	80.42
21-B2F	NN	89.88	A	9285	82.16
21-B3F	NN	49.14	A	10663	83.77
22-B1F	NH	46.79	A	10264	85.83
22-B2F	DH	19.08	A	14523	82.39
23-B3F	DH	59.35	A	5548	81.65
24-B1F	DH	19.75	A	12012	83.70
24-B2F	DD	62.52	A	3249	75.62
24-B3F	DD	3.60	A	7171	80.83

25-B1F	NH	63.67	A	10637	81.66
26-B2F	NN	32.98	A	7388	83.05
27-B2F	NH	57.57	A	11992	91.12
27-B3F	DD	45.78	A	12970	88.66
29-B1F	NH	55.69	A	16608	84.97
30-B2F	DD	41.33	A	8130	83.89
33-B2F	DH	2.99	A	7908	90.59
33-B3F	DH	28.31	A	7948	81.35
02-B1F	NH	58.59	B	7773	96.57
03-B3F	NH	64.97	B	14847	99.51
06-B3F	DH	3.59	B	9042	98.83
07-B2F	DD	2.82	B	3941	97.64
08-B1F	DD	6.08	B	16447	98.61
08-B2F	NH	3.44	B	12825	99.27
09-B2F	NH	5.80	B	24288	96.83
10-B2F	DH	3.55	B	20371	99.36
13-B1F	DH	2.93	B	7346	94.80
16-B1F	DH	6.29	B	13921	98.99
18-B1F	DH	1.57	B	10261	94.29
19-B2F	NN	3.09	B	16329	98.51
20-B2F	NH	57.16	B	13028	98.53
20-B3F	NH	7.57	B	14615	98.47
21-B1F	NN	2.85	B	6503	97.91
22-B3F	DD	3.01	B	9433	99.14
23-B1F	DH	2.41	B	11224	97.46
25-B2F	DH	2.65	B	19484	99.12
25-B3F	DD	70.32	B	15942	99.43
26-B1F	NN	2.85	B	11586	99.07
26-B3F	NN	3.55	B	14829	98.92
28-B3F	NN	3.71	B	17491	99.40
29-B3F	DH	2.71	B	7009	97.73
30-B3F	DH	3.12	B	28573	99.30
31-B1F	DD	7.26	B	17016	96.21
31-B3F	NH	2.50	B	12031	99.13
32-B1F	NN	4.77	B	13956	98.64

Table S3 Summary of ANOVA test statistics on microbiota alpha diversity (inverse Simpson index). All lakes refer to the analysis of effect of host species (dwarf and normal), lake (Cliff, East, Indian, Témiscouata and Webster). Second, the effect of host species is treated for each lake separately. Third, CAPTIVE refers to the analysis of effect of host group (dwarf, normal, hybrids D♀N♂ and N♀D♂), diet (Artemia only and Artemia with dry food) on all captive fish. Fourth, effect of captivity (wild and captive), dwarf only and normal only. F-Value is the value of the F statistic.

Fish group	Source of variation	Degrees of freedom	F value	P value
WILD				
	Species	1	0.439	0.510
All lakes	Lakes	4	2.304	0.064
	Lakes:Species	4	1.152	0.337
Cliff	Species	1	0.109	0.744
Est	Species	1	0.025	0.876
Indian	Species	1	2.026	0.169
Témiscouata	Species	1	1.557	0.225
Webster	Species	1	0.824	0.380
CAPTIVE				
	Group	3	0.599	0.620
	Diet	1	40.471	0.001
	Species:Diet	3	1.930	0.134
BOTH				
All fish	Captivity	1	8.915	0.003
Dwarf	Captivity	1	2.044	0.157
Normal	Captivity	1	2.040	0.157

Table S4 Diet influence on mass and length of captive whitefish sample. DD: dwarf whitefish, NN: normal whitefish, DH: hybrid F1 D♀*N♂, NH: F1 N♀*D♂. The diet group A is composed of Artemia and dry food; B is composed of Artemia.

Sample	Whitefish species	Tank	Diet Group	Body mass (gram)	Body length (cm)	Mass/Lenght
04-B2F	NH	B2	A	62.50	17.4	3.591
04-B3F	NH	B3	A	68.45	17.3	3.956
05-B1F	NN	B1	A	54.11	15.3	3.536
05-B2F	DD	B2	A	48.75	16.8	2.901
05-B3F	NH	B3	A	49.62	16.4	3.025
06-B1F	DH	B1	A	20.27	11.9	1.703
06-B2F	NH	B2	A	53.55	16.2	3.305
07-B1F	DD	B1	A	70.76	19.7	3.591
07-B3F	NN	B3	A	66.45	17.1	3.885
08-B3F	NH	B3	A	38.24	14.7	2.601
09-B1F	DD	B1	A	35.43	15.4	2.300
09-B3F	NN	B3	A	78.10	16.6	4.704
10-B1F	DD	B1	A	66.88	17.8	3.757
10-B3F	DD	B3	A	40.19	16.1	2.496
11-B1F	DD	B1	A	70.33	18.8	3.740
11-B3F	DD	B3	A	15.23	11.9	1.279
12-B1F	NH	B1	A	41.27	15.0	2.751
12-B2F	DH	B2	A	12.24	11.0	1.112
12-B3F	NH	B3	A	96.84	19.6	4.940
13-B2F	DH	B2	A	15.65	11.7	1.337
14-B1F	NH	B1	A	45.23	15.9	2.844
14-B3F	DD	B3	A	61.29	18.0	3.405
15-B1F	DH	B1	A	26.03	13.0	2.002
15-B2F	NN	B2	A	97.68	18.7	5.223
15-B3F	DH	B3	A	51.04	16.4	3.112
16-B3F	NH	B3	A	70.71	17.5	4.040
17-B1F	NN	B1	A	64.96	16.9	3.843
17-B2F	NN	B2	A	39.58	14.9	2.656
17-B3F	DD	B3	A	86.65	-	-
18-B2F	DD	B2	A	78.20	18.8	4.159
19-B3F	NN	B3	A	83.91	17.8	4.714
20-B1F	DD	B1	A	22.63	13.2	1.714
21-B2F	NN	B2	A	89.88	18.1	4.965
21-B3F	NN	B3	A	49.14	15.2	3.232
22-B1F	NH	B1	A	46.79	15.6	2.999
22-B2F	DH	B2	A	19.08	12.8	1.490
23-B3F	DH	B3	A	59.35	17.2	3.450
24-B1F	DH	B1	A	19.75	12.5	1.580
24-B2F	DD	B2	A	62.52	17.8	3.512

24-B3F	DD	B3	A	3.60	8.2	0.439
25-B1F	NH	B1	A	63.67	17.7	3.597
26-B2F	NN	B2	A	32.98	13.7	2.407
27-B2F	NH	B2	A	57.57	17.0	3.386
27-B3F	DD	B3	A	45.78	15.8	2.897
29-B1F	NH	B1	A	55.69	17.1	3.256
30-B2F	DD	B2	A	41.33	16.4	2.520
33-B2F	DH	B2	A	2.99	7.4	0.404
33-B3F	DH	B3	A	28.31	12.1	2.339
02-B1F	NH	B1	B	58.59	16.5	3.550
03-B3F	NH	B3	B	64.97	18.1	3.589
06-B3F	DH	B3	B	3.59	8.2	0.437
07-B2F	DD	B2	B	2.82	7.5	0.376
08-B1F	DD	B1	B	6.08	9.8	0.620
08-B2F	NH	B2	B	3.44	8.0	0.43
09-B2F	NH	B2	B	5.80	9.0	0.644
10-B2F	DH	B2	B	3.55	7.7	0.461
13-B1F	DH	B1	B	2.93	7.2	0.406
16-B1F	DH	B1	B	6.29	9.3	0.676
18-B1F	DH	B1	B	1.57	6.2	0.253
19-B2F	NN	B2	B	3.09	7.5	0.412
20-B2F	NH	B2	B	57.16	16.8	3.402
20-B3F	NH	B3	B	7.57	9.4	0.805
21-B1F	NN	B1	B	2.85	7.1	0.401
22-B3F	DD	B3	B	3.01	7.6	0.396
23-B1F	DH	B1	B	2.41	6.9	0.349
25-B2F	DH	B2	B	2.65	6.9	0.384
25-B3F	DD	B3	B	70.32	18.7	3.760
26-B1F	NN	B1	B	2.85	7.0	0.407
26-B3F	NN	B3	B	3.55	7.1	0.500
28-B3F	NN	B3	B	3.71	7.5	0.494
29-B3F	DH	B3	B	2.71	6.8	0.398
30-B3F	DH	B3	B	3.12	7.4	0.421
31-B1F	DD	B1	B	7.26	10.2	0.711
31-B3F	NH	B3	B	2.50	6.8	0.367
32-B1F	NN	B1	B	4.77	8.1	0.588

Table S5: Four Metastats tables with details of one-species-specific genera. DD: dwarf whitefish, NN: normal whitefish, DH: hybrid F1 D♀*N♂, NH: F1 N♀*D♂. The diet group A is composed of Artemia and dry food; B is composed of Artemia.

Metastats comparison DD and DH captive whitefish		
Genera	Diet group	Groups
Aquicella	B	DD
Flavobacterium	B	DD
Fusobacterium	B	DD
Lactobacillus	B	DD
Leuconostoc	B	DD
Sphingobium	B	DD
Streptococcus	B	DD
Weissella	B	DD
Aeromonas	A	DD
Kocuria	A	DD
Nocardioides	A	DD
Rhodobacter	A	DD
Corynebacterium	A	DH
Delftia	A	DH
Labrenzia	A	DH
Legionella	A	DH
Paracoccus	A	DH
Planctomyces	A	DH
Pseudomonas	A	DH
Stenotrophomonas	A	DH

Metastats comparison NN and DD captive whitefish		
Genera	Diet group	Groups
Acinetobacter	B	DD
Actinomyces	B	DD
Anoxybacillus	B	DD
Aquicella	B	DD
Barnesiella	B	DD
Brevundimonas	B	DD
Campylobacter	B	DD
Cerasibacillus	B	DD
Faecalibacterium	B	DD
Oceanobacillus	B	DD
Ohtaekwangia	B	DD
Prevotella	B	DD
Pseudolabrys	B	DD
Rhodococcus	B	DD
Sphingobium	B	DD
Sphingomonas	B	DD

Thermoflavimicrobium	B	DD
Vogesella	B	DD
Gp6	A	DD
Lactobacillus	A	DD
Streptococcus	A	DD
Alishewanella	B	NN
Aneurinibacillus	B	NN
Aquabacterium	B	NN
Arcobacter	B	NN
Arthrobacter	B	NN
Bifidobacterium	B	NN
Cellvibrio	B	NN
Corynebacterium	B	NN
Devosia	B	NN
Duganella	B	NN
Ethanoligenens	B	NN
Gemella	B	NN
Haemophilus	B	NN
Hyphomicrobium	B	NN
Megasphaera	B	NN
Methyloversatilis	B	NN
Novosphingobium	B	NN
Photobacterium	B	NN
Prauserella	B	NN
Pseudomonas	B	NN
Rothia	B	NN
Shewanella	B	NN
Sphaerobacter	B	NN
Tissierella	B	NN
Undibacterium	B	NN
Labrenzia	A	NN
Planctomyces	A	NN

Metastats comparison NH and DH captive whitefish

Genera	Diet group	Groups
Aliivibrio	B	DH
Aneurinibacillus	B	DH
Arcobacter	B	DH
Clostridium_sensu_stricto	B	DH
Clostridium_XI	B	DH
Comamonas	B	DH
Deinococcus	B	DH
Dickeya	B	DH
Dokdonella	B	DH
Elizabethkingia	B	DH
Enhydrobacter	B	DH
Exiguobacterium	B	DH

Gemella	B	DH
Legionella	B	DH
Limnohabitans	B	DH
Megasphaera	B	DH
Novosphingobium	B	DH
Phenylobacterium	B	DH
Proteiniclasticum	B	DH
Pseudomonas	B	DH
Rhodobacter	B	DH
Salinimicrobium	B	DH
Shewanella	B	DH
Sphingobacterium	B	DH
Sphingobium	B	DH
Thermomonas	B	DH
Wautersiella	B	DH
Aerococcus	A	DH
Alkanindiges	A	DH
Aquabacterium	A	DH
Delftia	A	DH
Desulfovibrio	A	DH
Flavobacterium	A	DH
Labrenzia	A	DH
Legionella	A	DH
Psychrobacter	A	DH
Stenotrophomonas	A	DH
Listonella	A	DH
Trichococcus	A	DH
Methylobacterium	A	DH
Paracoccus	A	DH
Aerococcus	B	NH
Leptotrichia	B	NH
Pediococcus	B	NH
Photobacterium	B	NH
Turneriella	B	NH
Aquitalea	A	NH
Bifidobacterium	A	NH
Brevundimonas	A	NH
Devosia	A	NH
Gemmata	A	NH
Hyphomicrobium	A	NH
Lactobacillus	A	NH
Oerskovia	A	NH
Ohtaekwangia	A	NH
Olsenella	A	NH
Pasteuria	A	NH

Metastats comparison NN and NH captive whitefish		
Genera	Diet group	Groups
Acinetobacter	B	NH
Aerococcus	B	NH
Brachymonas	B	NH
Brevibacterium	B	NH
Catonella	B	NH
Cupriavidus	B	NH
Dermacoccus	B	NH
Escherichia_Shigella	B	NH
Kocuria	B	NH
Leptotrichia	B	NH
Mycobacterium	B	NH
Neisseria	B	NH
Oceanobacillus	B	NH
Paenibacillus	B	NH
Pediococcus	B	NH
Peptococcus	B	NH
Rhizobium	B	NH
Roseomonas	B	NH
Sphingobacterium	B	NH
Turneriella	B	NH
Vibrio	B	NH
Acinetobacter	A	NH
Bifidobacterium	A	NH
Brevundimonas	A	NH
Chryseobacterium	A	NH
Collimonas	A	NH
Lactobacillus	A	NH
Ohtaekwangia	A	NH
Olsenella	A	NH
Streptococcus	A	NH
Weissella	A	NH
Acetobacterium	B	NN
Acidovorax	B	NN
Aeriscardovia	B	NN
Aneurinibacillus	B	NN
Aquabacterium	B	NN
Arcobacter	B	NN
Arthrobacter	B	NN
Bacillus	B	NN
Bifidobacterium	B	NN
Brevibacillus	B	NN
Cellvibrio	B	NN
Devosia	B	NN
Duganella	B	NN

Enhydrobacter	B	NN
Ethanoligenens	B	NN
Gemella	B	NN
Haemophilus	B	NN
Hyphomicrobium	B	NN
Legionella	B	NN
Massilia	B	NN
Megasphaera	B	NN
Methyloversatilis	B	NN
Novosphingobium	B	NN
Prauserella	B	NN
Pseudomonas	B	NN
Psychrilyobacter	B	NN
Rothia	B	NN
Sphaerobacter	B	NN
Sphingobium	B	NN
Sphingomonas	B	NN
Undibacterium	B	NN
Comamonas	A	NN
Elizabethkingia	A	NN
Flavobacterium	A	NN
Ilumatobacter	A	NN
Labrenzia	A	NN
Leuconostoc	A	NN
Listonella	A	NN
Psychrobacter	A	NN
Rhizobium	A	NN
Rhodopirellula	A	NN
Shewanella	A	NN