An Environmental DNA Metabarcoding Approach to Monitoring Red Algal Biodiversity

by

Josephine Crichton

BSc w/ Specialization in Ecology, Evolution and Environmental Biology, University of Alberta, 2020

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

Master of Science

in the Graduate Academic Unit of Biology

Supervisors: Gary W. Saunders, PhD, Biology

Adrian Reyes-Prieto, PhD, Biology

Examining Board: Michael Duffy, PhD, Biology, Chair
Jason Addison, PhD, Biology
Audrey Limoges, PhD, Earth Sciences

This thesis is accepted by the
Dean of Graduate Studies

THE UNIVERSITY OF NEW BRUNSWICK

October 2023

ABSTRACT

Biodiversity monitoring is an important tool in ecosystem conservation as it can provide reliable detection of invasive species, indicate resilience after disaster, and serves as evidence for successful management of resources or ecosystems. Despite biodiversity loss being one of our most critical conservation challenges, biodiversity monitoring in marine ecosystems is typically conducted using molecular-assisted alpha taxonomic (MAAT) surveys, which have several limitations. This study compares red algal species richness and compositional diversity detected using environmental DNA (eDNA) metabarcoding to that detected using a typical MAAT survey. My aim was to determine if eDNA analyses provide a cost-effective and comprehensive alternative to traditional survey methods, and to improve upon the currently available reference libraries for red algae. The results of this support a growing bank of literature suggesting that a combination of MAAT and eDNA methods remains the best approach for maximizing species detection.

DEDICATION

To my biggest supporters, my mom, Carey, and Michael.

ACKNOWLEDGEMENTS

I want to firstly acknowledge Dr. Gary Saunders for his continued passion and concern for my project. It is instructors, mentors, and scientists like you who get people excited about this field. I appreciate the candid conversations we have had over the last couple of years and the feedback and advice that I will take into my career and life moving forward. I will always think fondly of the field trips, except for the one when I got extremely seasick. Otherwise, I've had a great time. Thank you for being engaged, supportive, and keeping me on track. I want to equally thank Dr. Adrian Reyes-Prieto for your support during this project, and for answering my questions with patience and grace. I deeply appreciate your willingness to sit beside me at the computer, and really take the time to figure things out with me. I am grateful to have had two great supervisors throughout this degree. Thank you to Rene Malenfant for your direction and kindness, and for meeting with me even while on sabbatical.

I want to acknowledge the Department of Fisheries and Oceans for supporting and funding this project. This project would not have been possible without this support. I want to also acknowledge this work was conducted on the traditional land of the Passamaquoddy, Wolastoqiyik and Mi'kmaq people.

A big thank-you goes to my lab mates, Marie Dankworth, Josh Evans, Cody Brooks, Maggie Cassidy, Tanya Moore. From helping work through ideas, to teaching me how to run a PCR properly, to coffee breaks and early field days, you have all been incredibly supportive and helpful – thank you. I want to especially acknowledge Marie for being a huge source of inspiration throughout my time here, and for grounding me

when I needed it most. Our time together will always hold a special place in my heart, and I will miss seeing you every day.

Cody, thank you for being patient with me in the very early stages of my degree. I distinctly remember having you watch each and every step of a PCR I was running to ensure I was doing it properly. Josh, you have similarly offered me patience and honest, critical thought about my project. I'm sorry for keeping you in the hallways chatting for so long (not really). Maggie, our trip to NEAS together is a memory I will not forget – nothing bonds two new students like an 8-hour drive with semi-strangers. I'm glad I had you for that trip and most of this degree. Thank you for always celebrating the little wins with me and always being willing to come sample, stare at Geneious, or get a coffee with me.

Thank you to the fellow researchers and staff at UNB for their help and inspiration. To Cecilio Valadez-Cano, I really can't thank you enough for your help writing a script that I wound up using nearly every day for months. To Shawn MacLellan, for welcoming me to come sit in your office and talk through things. To Dr. Katy Hind, for sparking my love of seaweed back in 2019 in Bamfield. For introducing me to her PhD supervisor, Gary, who would soon become my supervisor as well. For showing me that you can be both a scientist and a mother.

Thanks to Craig and Ginny for the lab T-shirt. Thank you for thinking of me and for the all the stories about Gary. Similarly, thank you to Amanda Savoie, Lesleigh Kraft, Kirby Morrill and Trevor Bringloe for the stories and advice on getting through grad school. I think it mostly helped. It has been a pleasure getting to know all of you. Thank

you to Joe George for making americanos out on the boat and for making first-aid kind of fun.

Thank you to Margaret and Hannah and the staff at the Grad House, for being great baristas and friends, and providing me a place to write nearly this entire thesis. Thank you to my book club friends for reading drafts of my thesis, even if it was mostly as a joke. Conor, Bradley, Anna, Grace, Jessie, Danny, Jon, Vish, Ben Saad, Ben Dow, and Brayton: you guys made the last year so much fun. I'm grateful to have the unwavering support of my friends, family, and partner, despite never really grasping this whole ordeal. Thank you to Michael for believing in me and for all of your support.

Table of Contents

List of Tables

List of Figures

List of Symbols, Nomenclature or Abbreviations

Abbreviation	Description
BLAST	Basic Local Alignment Search Tool
BOLD	Barcode of Life Database
CCS	Circular Consensus Sequence
CEMAR	Centre for Environmental and Molecular Algal Research
COI-5P	Cytochrome c Oxidase Subunit 1 (five prime portion)
eDNA	Environmental DNA
Ga	Billion years
IMR	Integrated Microbiome Resource
MAAT	Molecular-assisted Alpha Taxonomy
PR^2	Protist Ribosomal Reference
PR^2 RED1	Protists Ribosomal Reference with additional red barcodes
*rbc*L	Ribulose-1,5-Bisphosphate Carboxylase Large Subunit gene
SCUBA	Self-contained Underwater Breathing Apparatus
SMRT	Single-molecule Real-time
SSU rRNA	Small Subunit Ribosomal Ribonucleic Acid
UNB	University of New Brunswick

Introduction

Marine biodiversity is experiencing large-scale changes on a global level (Craig 2012, Worm & Lotze 2021), making reliable biodiversity monitoring a particularly topical challenge. Biodiversity monitoring aids in the conservation of economically and ecologically important species by providing a baseline of species composition and evidence for successful management practices (Gold et al. 2021, Lindenmayer et al. 2012). Biodiversity data can inform policymakers to ensure marine ecosystems are being sufficiently protected and monitored regularly (Niemelä 2000). The monitoring of red algae is becoming increasingly critical as their importance in the food, medicine, and textile industries continues to grow (Abdelgawad et al. 2020, Wang et al. 2020). Red algae are therefore inextricably linked with not only the health of marine ecosystems but with human health and the economy as well. As red algae become more popular in large-scale industry, the threat of biodiversity loss and consequently reduced ecosystem resilience grows due to overexploitation, urbanization, habitat destruction, and invasive harvesting practices. Thus, finding non-invasive and accessible methods to monitor species of interest is a large part of the biodiversity and conservation challenge.

Biodiversity Monitoring as a Conservation Tool

The objectives of biodiversity monitoring are to provide information on trends in population changes, to detect invasive species, to monitor endangered species, to set priorities for conservation, and to inform ecosystem management on where conservation efforts have been successful (Barnes et al. 2014, Lindenmayer et al. 2012, Niemela 2000). Understanding the current state of biodiversity in our marine ecosystems allows us

to monitor the impacts of harvesting, climate change, urbanization, pollution, and other ecosystem change over time. Species richness and biodiversity enhance ecosystem productivity and stability and are essential for the maintenance of productive ecosystem goods and services (Gamfeldt et al. 2015, Loreau et al. 2001).

Biodiversity monitoring can detect the early presence of invasive species, many of which can increase biotic homogenization (Muthukrishnan & Larkin 2020). When implemented regularly, biodiversity surveys can be directly compared to assess biotic homogenization between previously geographically distinct communities. Biotic homogenization often results in reduced ecosystem function and services, including the provision of food and habitat for the fish and invertebrates we consume, pharmaceuticals, maintenance of water quality, protection from storms, and carbon capture (FAO 2019, Worm et al. 2006).

Invasive species are a threat to native red algae in the Bay of Fundy as demonstrated with the introduction of *Bonnemaisonia hamifera* Hariot (Nylund et al. 2013), which now occupies all Atlantic Canadian waters excluding the St. Lawrence estuary (Natural Resources Canada, 2002). The weedy nature and mat-like growth of this species has allowed it to monopolize some subtidal locations in the Bay of Fundy forming extensive mats formerly occupied by native species (Saunders, unpublished data). Marine algae can be transported across oceans in bilge water or attached to ships (Saunders & Withall 2006), and marine ecosystems are therefore at high risk of invasion by alien species. Many invasive species have broad physiological tolerances and may be outcompeting native red algae, introducing disease, and altering community structure (Mathieson et al. 2008, Newton et al. 2013). Successful establishment of invasive species

can have dramatic impacts by shading understory species, monopolizing space required by niche species, and altering the composition of food ingested by grazers (Bax et al. 2003, Mineur et al. 2015). The interactions among native and introduced species are not all inherently negative, as some introduced species have increased productivity and can provide habitat and food for local fish and invertebrates (Mineur et al. 2015), but any changes to community composition should be monitored to detect any long-term effects of introduced species. Our first line of defense against invasive species is their detection; thus, finding cost-effective and reliable methods of detection is key to the successful management of invasive species and subsequent biotic homogenization and potentially harmful ecosystem changes.

Despite comprehensive biodiversity data that exist in some regions for certain groups of interest, temporal and spatial limitations make it hard to discern large-scale trends in biodiversity. Without this bigger picture, we cannot inform policy and make realistic goals towards conservation and resource/ecosystem management (Lindenmayer et al. 2012). Implementation of effective standard operating procedures for marine ecosystems requires robust and long-term biodiversity data to plan and prioritize accordingly (FAO 2019).

Study System

Our study focuses on the red algae of the Bay of Fundy, in New Brunswick, Canada. The Bay of Fundy is an inlet of the Atlantic Ocean located in the Gulf of Maine between New Brunswick and Nova Scotia (Figure 1.1). Our study focuses on two sites on the New Brunswick side of the Bay of Fundy, herein referred to as the Lepreau site and

the Wallace Cove site (Figure 1.1). The Bay of Fundy is home to the highest tides in the world, which have strong impacts on the biological, chemical, and sedimentological processes in the coastal regions of New Brunswick (Bates et al. 2001, Desplanque & Mossman 2001). This unique and productive ecosystem has been used for aquaculture, commercial fishing, and tidal energy sites. The aquaculture and fishing industries contribute over 3,500 jobs, and over $30 million in wages in the province (Dept. of Agriculture, Fisheries & Aquaculture 2023). This ecosystem is integral to the livelihood of the province, though baseline biodiversity surveys for many species, including red algae, are rare (Bates et al. 2001, Kalu et al. 2023), making it difficult to assess the impacts of this continuing anthropological development in the Bay of Fundy.

Red algae are one of the oldest extant eukaryotic lineages on Earth, the oldest red algal fossil estimated to be 1.2–1.1 Ga (Yang et al. 2016, Yoon et al. 2017). Thus, they have deeply rooted evolutionary relationships with other marine species (Minhas et al. 2020). They occupy nearly all marine environments on Earth and are integrated into many different ecosystems, including some freshwater environments (Szinte et al. 2020). Due to their prevalence in marine and aquatic ecosystems, research is starting to use red algal properties, such as thallus yellowing and depth distribution, as indicators for ecosystem health (Chankaew et al. 2015, Kobayashi & Fujita 2014, Rinne & Kostamo 2022). Despite their obvious presence in the Bay of Fundy, the monitoring of this group is irregular and rarely considered in impact assessments and management plans (Norton et al. 1996).

Red algae often exhibit phenotypic plasticity, making them hard to identify and distinguish morphologically; in addition, they can exhibit a heteromorphic alternation of

generations in which the gametophytic and sporophytic stages of the same species exhibit completely different morphologies (Saunders 2005). Historical records of these species can be limited to the most easily found or identified stage, and many species are still being discovered with the assistance of molecular data. Strictly morphological surveys for this group can be unreliable and have long caused confusion among algal taxonomists (Saunders 2008), requiring the intervention of molecular survey methodologies. These complexities in our study system have led to a gap in the monitoring of red algae in the Bay of Fundy, which is required to inform comprehensive ecosystem management and conservation practices.

Conducting Marine Biodiversity Surveys

Current seaweed monitoring is completely based on morphology, or a combination of morphological observations and DNA barcoding in an approach referred to as molecular-assisted alpha taxonomy (MAAT) (Saunders 2005). DNA barcoding refers to sequencing a gene fragment from a previously identified specimen to form a database against which we can compare unknown specimens (Moritz & Cicero 2004, Prasanthi et al. 2020). MAAT surveys provide robust and reliable identification and have the additional benefit of a physical specimen that can be analyzed microscopically to confirm or learn microscopic characters of the taxa. Despite the desirable controllability of MAAT surveys, many limitations remain. MAAT surveys depend on extensive taxonomic expertise, long field hours, SCUBA equipment and personnel, weather-dependent fieldwork, the actual harvest of plants, and the development of taxa-specific primer sets (Knudsen et al. 2022). Another concern with MAAT surveys is observer bias,

which is affected by dive conditions and taxonomic expertise. Red algae are notoriously difficult to identify, making reliable identification based on morphology a demanding task, even to family or genus level for some groups. This compounds the problem of the continuing decline of taxonomic expertise across nearly all study organisms (Agnarsson & Kuntner 2007). These factors influence the scope and frequency at which biodiversity monitoring can be reliably conducted, thus reducing the accessibility and comprehensiveness of biodiversity data. These limitations have resulted in a large-scale movement towards metabarcoding methods, in which multiple species can be identified from a single environmental sample.

Research Approach

To address these limitations, we trialed an environmental DNA (eDNA) metabarcoding approach using the widely used small subunit (SSU) rRNA gene to survey red algal biodiversity. Due to the numerous complexities of red algal sampling, this group offers a unique opportunity to test both the utility and limitation of eDNA metabarcoding. The eDNA methodology is an already established method in microbial biology and is growing in popularity amongst aquatic and marine biologists (Bohmann et al. 2014, Deiner et al. 2017). eDNA surveys provide a less invasive alternative to more traditional sampling methods, require no taxonomic expertise to collect, and are not reliant on SCUBA divers. eDNA sample collection is safer, faster, and more cost-efficient than traditional sample collection. The accessibility of this method is an obvious incentive for laboratories and funding agencies that currently spend considerable time and money on fieldwork and sample collection.

Beyond the reduced cost, eDNA has proven to be highly sensitive and can detect species frequently overlooked by visual surveys (Port et al. 2016). Yamamoto et al. (2017) found that an eDNA method found 40% more fish species than 14 years of traditional visual surveys. This presents a potentially significant advantage for phycologists as many marine algae are difficult to identify, and experts in the field are few – meaning some species are routinely misidentified or overlooked when conducting visual floristic surveys (Saunders 2008). Additionally, rare species are best sampled with methods that minimize physical harm and destruction to the organisms, their habitat, and their community (Pikitch 2018).

Environmental DNA generally refers to DNA retained from an environmental sample (water, air, soil, tissue, etc.) (Barnes & Turner 2016, Taberlet et al. 2012). Algal DNA sources are released into the water by means of detritus and reproductive spores and gametes. Preliminary studies have begun to estimate the persistence of eDNA in a marine environment, but many unknowns remain about the spatial and temporal scopes of eDNA, which depend on hydrodynamics and abiotic factors (Collins et al. 2018, Lacoursiere-Roussel et al. 2018). eDNA from killer whales has been found up to two hours after the known passage of the whale, suggesting that eDNA may be quite persistent even in a turbulent marine environment (Baker et al. 2018). However, some studies have shown the inability to detect target species even in known frequented areas (Rees et al. 2014, Thomsen et al. 2012). This suggests that eDNA persistence and detectability depends on the system and species in question. More research must be done to understand how far the DNA can travel from its source, how long it can remain stable, and the sampling effort required for optimal detection.

Many red algal species that are overlooked or misidentified in the field are epiphytic, small, or cryptic. Aiming to capture microscopic and epiphytic fragments, we used an intertidal rock scrape method. This method is expected to be efficient in capturing the many epiphytes present on marine algae, the species that grow clumped together or on top of one another, and the gametes, spores and detritus left on the algae and rocks.

DNA can be extracted from the rock scrape samples using a variety of DNA extraction techniques for subsequent sequencing. The two widely used sequencing systems in metabarcoding are Illumina (short reads; 150–350 bp) and Pacific BioSciences (PacBio; long reads; 350 bp – 25 kb) technologies. The selection of the sequencing platform depends on the desired sequence length of the molecular marker. Accurate long-read sequencing (e.g., PacBio) is a relatively new method that has allowed for many advancements in variant detection and metagenomics (Amarasinghe et al. 2020). Because PacBio generates longer sequences than the widely used Illumina technology, it can improve taxonomic resolution with comparable accuracy. The PacBio technology produces highly accurate high-fidelity (HiFi) reads using circular consensus sequencing of up to 25,000 bases at >99% accuracy (Hon et al. 2020). This is the platform we chose to sequence our SSU amplicons from rock scrape samples to target both gene length and sequencing accuracy. This study is the first to use long-read sequencing technology to investigate red algal biodiversity.

Bioinformatic analyses of metabarcoding data require the generation of an accurate and comprehensive taxonomic reference library of the selected molecular marker, and rapid, accurate pipelines to analyze large amounts of sequence data. Public

databases such as GenBank (Benson et al. 2013) span a large breadth of genetic markers and species diversity. Though public databases typically have some basic quality checks in place to minimize flawed sequence data may arise from the misidentification of vouchered material, the amplification of a symbiont or endoparasite, lab contamination, PCR error, generation of chimeric sequences, or simply human error in data entry (Leray et al. 2019, Meiklejohn et al. 2019). In this study, we supplemented the public SSU Protist Ribosomal Reference (PR^2) version 4.14.0 (https://github.com/pr2database/pr2database/releases/tag/v4.14.0) database with newly generated data for locally relevant species to capitalize on the breadth of the public database, while enhancing the reliability of sequences for regionally important species.

Our eDNA metabarcoding sequences were matched to sequences from our SSU supplemented reference database. We compared these results to species lists recorded by an expert in the field conducting a MAAT survey. We explored the overlap of species and genera detected with each survey type using a sequence similarity index and compared the species richness detected by each survey to examine the similarities and differences between the two methodologies.

My thesis aims to 1) enhance the completeness of an SSU Atlantic red algal reference library, and 2) compare red algal biodiversity results from eDNA surveys to those of MAAT surveys.

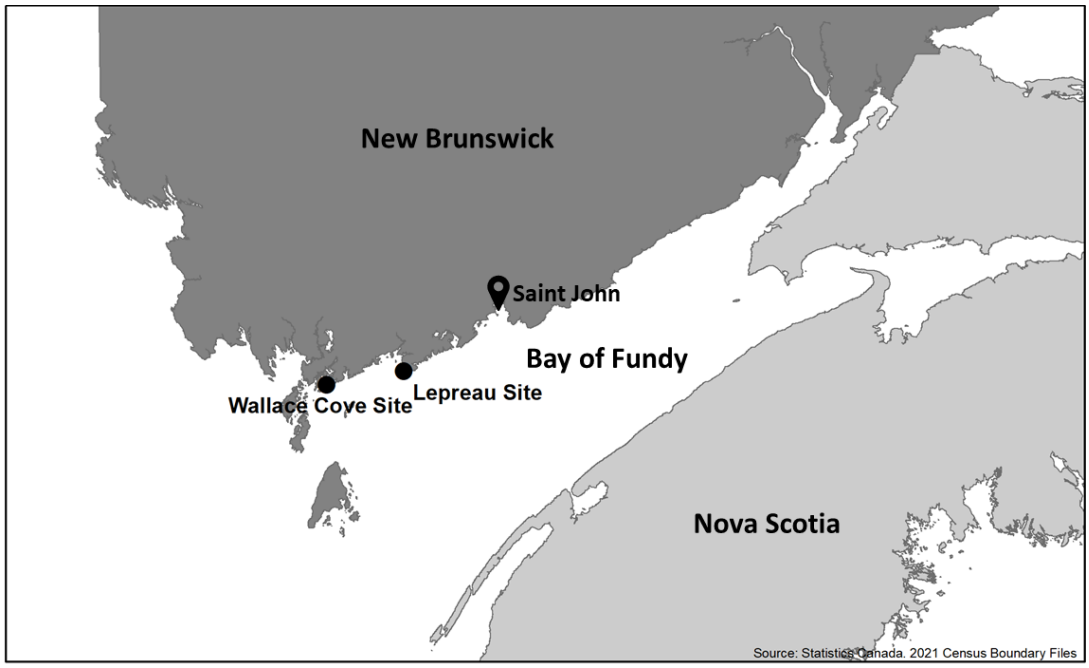

Figure 1.1 Map of both study sites (black circles) in the Bay of Fundy, New Brunswick.

Methods

MAAT Survey

MAAT surveys at each study site were completed by Dr. Gary Saunders, a taxonomic expert specializing in algal taxonomy, in the high to low intertidal zone by searching identifiable red algal species in the field and collecting species that were difficult to identify for subsequent microscopy and/or molecular identification. The Lepreau and Wallace Cove MAAT surveys were conducted May 29 and June 28, 2021,

respectively. The DNA from the collections was extracted according to the extraction protocols outlined in Saunders & McDevit (2012), and the *rbc*L-3P region was amplified using the primers and PCR settings outlined in Saunders & Moore (2013).

Reference Library Generation

We assembled a list of red algal species that have been collected from the Canadian Arctic and Atlantic regions to Rhode Island. We generated full-length SSU sequences for taxa without data currently available in GenBank for specimens with vouchered material in the Connell Memorial Herbarium (UNB, Fredericton) and associated DNA previously extracted with reported protocols (Saunders & McDevit 2012). The SSU region (1800 bp) was PCR amplified using external primers G01 (5'-CACCTGGTTGATCCTGCCAG-3'; forward) and G07 (5'-AGCTTGATCCTTCTGCAGGTTCACCTAC-3'; reverse) and sequenced with two external and two internal primers G04 (5'-CAGAGGTGAAATTCTTGGAT-3') and G14 (5'-CCTTGGCAGACGCTTTCGCAG-3') (Saunders & Moore 2013). The thermal profile for PCR amplification of the SSU includes an initial denaturation at 94°C, 38 cycles of 94°C for 30 s, 55°C annealing for 30 s, and 72°C extension for 90 s, and a 72°C final extension for 7 min (Saunders & Moore 2013). All PCR amplifications included a negative control with molecular grade water replacing template DNA. PCR products were run through a 0.8% agarose gel and stained with SYBR-safe dye to check for amplification and expected band size. Successful amplicons were sent for Sanger sequencing at Genome Quebec.

Sequences returned from Genome Quebec were edited (trimmed, forward and reverse reads aligned) using Geneious v 8.1.9 (www.geneious.com; Kearse et al. 2012). The new SSU barcode sequences were added to the public PR^2 database (Guillou et al. 2013) in FASTA format to be used as our improved reference library. These sequences can be found at https://github.com/josccrichton/PR2-Additions/tree/main. The PR^2 database with the additional red algal barcodes is herein referred to as PR^2 RED1.

eDNA Survey: Sample Collection and Processing

The Wallace Cove eDNA survey is made up of three sampling episodes from the spring (May 7, 2021), summer (July 9, 2021), and fall (September 8, 2021). The Lepreau eDNA survey is comprised of two sampling episodes, spring (May 7, 2021) and fall (September 8, 2021). We sampled both sites at low tide, and 10 scrapes were collected at random intervals along a 30 m transect line along the low intertidal. We used scissors to trim any large algae off the top of the sample and scrape the bottom 5 cm of wet turf closest to the rock using a fixed blade scraper; a 5 cm^2 area of material was scraped and placed in a Falcon™ tube filled with seawater on ice until the sample could be processed. The scraper was rinsed with freshwater between replicate scrapes at one location and washed with 70% ethanol between sites.

Samples were processed at the University of New Brunswick Centre for Environmental and Molecular Algal Research (CEMAR). Wet turf samples were drained and wrapped in paper towel and placed into silica. Samples were placed to dehydrate for two days; we then used liquid nitrogen to grind the samples into dry powder with 5 mg of the dry powder used for DNA extraction. The dry powder was used for DNA extraction

by hand using the ProMega PCR Preps DNA Purification System kit (protocol; Saunders & McDevit 2012). eDNA samples were quantified with the Qubit 2.0 fluorometer to ensure the DNA concentrations met the Integrated Microbiome Resource's (IMR) minimum requirements of >1ng/µL, preferably >10ng/µL. For quality assurance, we checked absorbance ratios (A260/230) of a subset of our samples using a ThermoScientific NanoDrop™ Spectrophotometer.

eDNA Metabarcoding: Sequencing

We sent 20µL of extracted DNA from each rock scrape (30 from Wallace Cove, 20 from Lepreau) for PacBio long-read sequencing at the IMR in Halifax, N.S. The IMR utilizes the PacBio Sequel 2 system and Single Molecule, Real-Time (SMRT) sequencing technology (Integrated Microbiome Resource, 2014). The SSU gene was amplified and sequenced using their standard full-length 18S fusion primers NSF4 (5'-CTGGTTGATYCTGCCAGT-3'; forward), and EukR (5'-TGATCCTTCTGCAGGTTCACCTAC-3'; reverse) which include the PacBio adapters, unique barcodes, and the specified gene region. Our samples were sequenced using the default parameters in the SMRT Link 10.2 to generate minimum 99% consensus (p. 76; SMRT Link User Guide 2021). The resulting circular consensus sequence (CCS) files were returned to us in FASTQ format. HiFi (highly accurate long sequencing) CCS reads have a Phred quality score >20 and are generated by creating a consensus of subreads sequenced from a single DNA molecule. Only HiFi CCS reads were considered in our analyses.

Bioinformatics

The initial analysis of the produced PacBio data included dereplication (identification of unique sequences), chimera detection (artifact sequences amplified from two different DNA templates during PCR), and taxonomic classification on the HiFi CCS reads (Table 1.1). First, we set minimum and maximum allowances for read length between 1000 and 2500 bp (SSU: 1800 bp) and dereplicated the sequences using *vsearch* (Rognes et al. 2016), retaining one copy of each sequence variant and keeping abundance counts. We removed chimeric sequences *de novo* also using *vsearch* (Rognes et al. 2016), as well as singletons and doubletons (i.e., sequences with an abundance of <3). The non-chimeric sequences were BLASTed (BLASTN; Camacho et al. 2009) against the PR^2 RED1 database, and all top hits with E-values and percent identity above our specified thresholds were retained (Table 1.1). The top BLASTN hit was then extracted according to percent sequence identity (min 99.5%) and highest query coverage and for each sequence variant calculated using a custom python script (https://github.com/josccrichton/PR2-Additions/tree/main). All software used and parameters and options for the dereplication, chimera detection, and taxonomic assignment steps are listed in Table 1.1. After the BLASTN step, all non-red algal matches were removed prior to final analyses.

Table 1.1. Programs and options selected for dereplication, chimera detection, and taxonomic assignment.

Program	Options
vsearch - dereplication	--derep_fulllength fasta.file --sizeout --minseqlength 100 --maxseqlength 2500 --relabel_sha1 --fasta_width 0
vsearch - sorting	--sortbysize fasta.file
vsearch - chimera detection	--uchime_denovo *fasta.file* --chimeras *output.file* --nonchimeras *output.file*
BLAST+ - database formatting	--makeblastdb *db.file* --dbtype nucl
BLAST+ - taxonomic assignment	--blastn --evalue 1e-15 --outfmt 6 --“qseqid sseqid pident length mismatch gaps qstart qend sstart sendevalue bitscore qlen” --perc_identity 99.5

Survey Comparisons

We selected two quantitative measures to compare the MAAT survey to the eDNA survey. First, we examined taxonomic richness detected using both survey types by determining the number of species (species richness) and genera detected by each method. Second, we evaluated the compositional similarity between the MAAT and eDNA surveys by calculating the Jaccard similarity coefficient in R Statistical Software (v3.4.4; R Core Team 2021) using the qvalue and jaccard packages (v0.1.0; Chung et al. 2018). The Jaccard index ranges from 0 to 1, 0 reflecting no common detections between the two groups, and 1 meaning the two lists are identical (Jaccard 1908). We did this analysis at both the species and the genus level. Additionally, we calculated the compositional similarity between the MAAT survey and one eDNA sampling episode to examine the impacts of seasonal species on our results.

Results

Reference Library Generation

We generated novel sequences for 84 species of red algae with NW Atlantic distribution using PCR amplification and Sanger sequencing (https://github.com/josccrichton/PR2-Additions/tree/main). The PR^2 RED1 library provides a total of 2,478 red algal SSU barcodes as reference, representing 1,064 red algal species.

eDNA Metabarcoding

Of the 50 eDNA samples from rock scrapes used for this study, eight samples sequenced poorly and were discarded, reducing our final number of samples to 42. We had 15 samples for the Lepreau site and 27 for the Wallace Cove site, generating a total of 920,658 raw long reads for both sites. The average number of reads per sample was 16,285. The maximum number of reads per sample was 77,694; the minimum was 1,197.

Wallace Cove Site

Of 334,020 reads, 4,970 sequences were discarded due to read length parameters. After dereplication, 157,183 unique sequences remained. We then removed singletons and doubletons resulting in 4,135 unique sequences. We discarded 765 chimeric sequences, which resulted in 3,388 unique non-chimeras from Wallace Cove. The similarity search identified 55.79% of the non-chimeras as red algal sequences, distributed in 19 species in 18 different genera. The remainder of the sequences were metazoans, dinoflagellates, oomycetes, ochrophytes, chlorophytes and ciliates. The seasonal breakdown of the read counts can be found in Table 1.2.

Lepreau Site

Of 586,638 reads, 3,800 sequences were discarded due to read length parameters. After read length adjustments and dereplication, 385,904 unique sequences remained. After the removal of singletons and doubletons, 6,602 unique sequences remained. We discarded 385 sequences as chimeric and kept 3,416 non-chimeras. Of the non-chimeras, 67.42% were red algal, representing 21 red algal species and 17 genera. The seasonal breakdown of read counts can be found in Table 1.2.

Table 1.2. Read counts at quality filtering steps split by sampling site and season.

Season and Site	Read Count	Uniques	Abundance ≥3	Chimeras	Non-chimeras
Spring Wallace Cove	44,582	20,311	577	34	543
Summer Wallace Cove	164,462	68,408	1,940	181	1,759
Fall Wallace Cove	124,976	68,464	1,636	550	1,086
Spring Lepreau	118,600	47,841	1,683	128	1,555
Fall Lepreau	157,460	72,172	2,118	257	1,861

Survey Comparisons

Wallace Cove Site

The MAAT survey at our Wallace Cove site detected 27 red algal species, outperforming the eDNA survey in terms of species richness, which detected 19 species (Table 1.3). The two surveys were 50% similar in species composition (Jaccard index = 0.500 with 15 common detections (Figure 1.2). Both eDNA and MAAT surveys made unique detections (species detections that were only made via one method). eDNA uniquely detected *Acrochaetium secundatum, Antithamnionella floccosa, Bonnemaisonia hamifera, and Polysiphonia* sp. 2 stricta. Between both eDNA and MAAT surveys, we detected 31 genetic groups at the Wallace Cove site. At the genus level, the MAAT survey detected 23 genera, where eDNA detected 18. The surveys were 57.6% similar in composition at the genus level (Jaccard index = 0.577), detecting 15 of the same genera.

Our MAAT survey at this site was conducted on June 28, 2021, so we also compared this survey to our summer eDNA sampling episode conducted on July 9, 2021. The summer eDNA sampling episode detected 12 genetic groups, missing *Bonnemaisonia hamifera, Clathromorphum circumscriptum, Corallina officinalis, Palmaria palmata, Plumaria plumosa,* and *Rhodochorton purpureum,* which were only found in either the spring or fall sampling episodes. The summer sampling episode and the MAAT survey performed 31% similarly (Jaccard index = 0.310), detecting nine of the same species. At the genus level, they were 34.6% similar in composition (Jaccard index = 0.346), detecting nine of the same genera.

Lepreau Site

The MAAT survey at the Lepreau site detected 27 red algal species, whereas eDNA detected 21 species (Table 1.3). The species compositions of the two surveys were 46.9% similar, (Jaccard index = 0.469) detecting 15 of the same species (Figure 1.3). Using both eDNA and MAAT surveys, we detected 32 genetic groups. At the genus level, the MAAT survey detected 24 genera, where eDNA detected 17. The surveys were 51.9% similar in composition at the genus level (Jaccard index = 0.519), detecting 14 of the same genera.

Our MAAT survey at the Lepreau site was conducted on May 29, 2021, and thus we also compared this survey to our spring eDNA sampling episode conducted on May 7, 2021. The MAAT survey detected a richness of 27, whereas the eDNA episode detected 19 species. The spring sampling episode and the MAAT survey were 40.6% similar in composition (Jaccard index = 0.406), detecting 13 of the same species. *Ahnfeltia plicata*

and *Phycodrys* sp. 1NB were only detected in the fall sampling episode. At the genus level, they were 48.1% similar in composition (Jaccard index = 0.481)

Table 1.3. Red algal species present using MAAT and eDNA survey methods at the Wallace Cove and Lepreau study sites.

	WALLACE COVE			*LEPREAU*		
Species Name	**MAAT**	**eDNA Complete**	**eDNA Summer**	**MAAT**	**eDNA Complete**	**eDNA Spring**
Acrochaetium secundatum	0	1	1	0	0	0
Ahnfeltia plicata	1	0	0	1	1	0
Antithamnionella floccosa	0	1	1	1	1	1
Bonnemaisonia hamifera	0	1	0	0	0	0
Ceramium deslongchampsii	1	0	0	0	0	0
Ceramium virgatum	1	1	1	1	1	1
Chondrus crispus	1	1	1	1	1	1
Clathromorphum circumscriptum[a]	1	1?	0	1	1?	1?
Clathromorphum sp. 1 circumscriptum[a]	0	1?	0	0	1?	1?
Coccotylus brodiei	0	0	0	0	1	1
Corallina officinalis	1	1	0	1	1	1
Corallina sp.	0	0	0	0	1	1
Cystoclonium purpureum	1	1	1	1	1	1
Devaleraea ramentacea	1	1	0	1	0	0

Dumontia contorta	1	1	1	1	0	0
Euthora cristata	1	0	0	1	0	0
Lithothamnion lemoineae	1	0	0	0	0	0
Mastocarpus stellatus	1	1	1	1	1	1
Membranoptera fabriciana	0	0	0	1	0	0
Neopyropia leucosticta	1	1	1	1	0	0
Palmaria palmata	1	1	0	1	0	0
Peyssonnelia rosenvingei	0	0	0	0	1	1
Phycodrys fimbriata	1	0	0	1	1	1
Phycodrys rubens	1	0	0	0	0	0
Phycodrys sp. 1 NB[b]	0	0	0	1	1	0
Phyllophora pseudoceranoides	0	0	0	1	0	0
Phymatolithon sp.3AT crust	0	0	0	0	1	1
Plumaria plumosa	1	1	0	1	1	1
Polyides rotunda	0	0	0	1	0	0
Polysiphonia sp. 2 stricta	0	1	1	1	1	1
Polysiphonia sp. 1 stricta[c]	1	1?	1?	1	1?	1?
Polysiphonia sp. 3 stricta[c]	1	1?	1?	0	1?	1?

Ptilota serrata	1	0	0	1	0	0
Rhodochorton purpureum	1	1	0	1	1	1
Rhodomela lycopodiodes	1	1	1	1	1	1
Rhodomela sp.	0	0	0	1	0	0
Rhodophysema georgei	1	0	0	0	0	0
Scagelia pylaisaei	1	0	0	0	0	0
Vertebrata lanosa	1	1	1	1	0	0
Vertebrata sp. 2fucoides	1	0	0	0	0	0
Wildemania amplissima	0	0	0	0	1	1
Wildemania miniata	1	0	0	1	1	1
TOTAL RICHNESS	**27**	**19**	**12**	**27**	**21**	**19**

[a] SSU is identical for *Clathromorphum circumscriptum* and *Clathromorphum* sp. 1circumscriptum; counted as a single eDNA detection.
[b] Reference sequences are inadequate to distinguish between *Phycodrys rubens* and *Phycodrys* sp. 1NB; eDNA match to *Phycodrys* was assigned to *Phycodrys* sp. 1NB using ecological context.
[c] Reference sequences inadequate to distinguish between *Polysiphonia* sp. 1stricta and *Polysiphonia* sp. 3stricta; counted as a single eDNA detection.

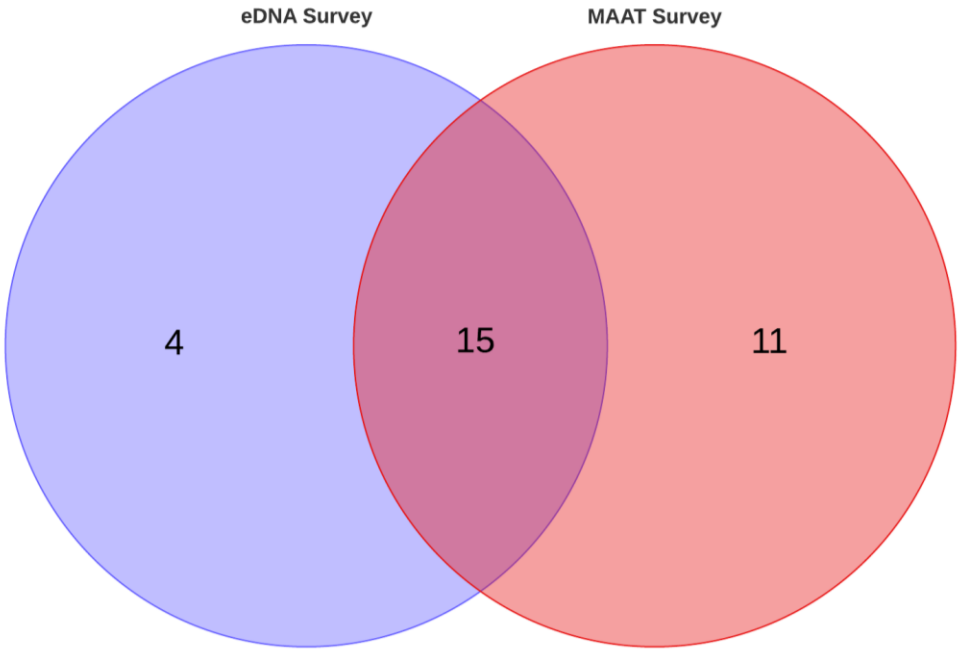

Figure 1.2. Venn diagram depicting the number of shared detections between the two survey types at the Wallace Cove site.

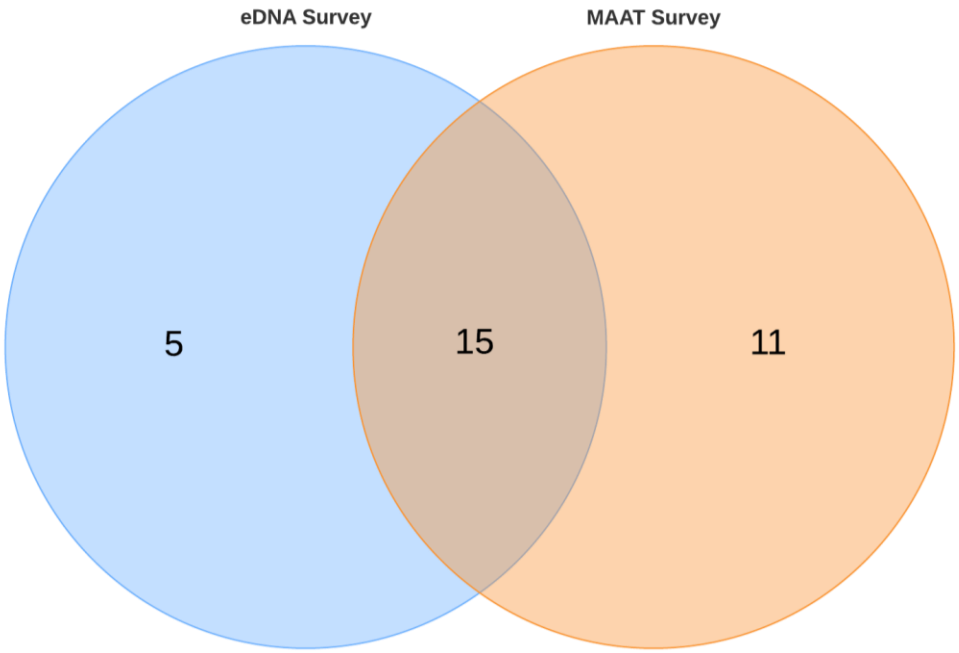

Figure 1.3. Venn diagram depicting the number of shared detections between MAAT and eDNA survey types at the Lepreau site**.**

Discussion

Importance of Taxonomists

Our results suggest a seasoned taxonomist can still outperform an eDNA survey in detecting red algal species richness in our study system. However, eDNA may provide a middle ground that allows non-experts to conduct surveys and perform better than they would have had they used a MAAT survey method. Additionally, taxonomists are required to verify the identity of voucher material used for reference sequences. In general, the field of taxonomy can and should work in harmony with new technologies and methodologies. Both survey types combined detected higher species richness than either method alone, and thus our results concur with those of other studies (Deiner et al. 2017, Grealy et al. 2015, Mas-Carrio et al. 2022, Milla et al. 2022), suggesting eDNA metabarcoding studies are a complement rather than a replacement for taxonomists and traditional surveys.

Comparison to MAAT Surveys

MAAT surveys require extensive taxonomic knowledge, extended hours in the field, dependence on weather conditions, and significant lab work to PCR amplify and sequence the collections to confirm their identity. These limitations contribute to a lack of knowledge surrounding our current state of biodiversity, which makes it difficult to monitor changes caused by climate change, invasive species, and other anthropological factors. This has inspired the development of new, rapid biodiversity-monitoring techniques that are relatively easy and inexpensive to implement. eDNA metabarcoding

offers an alternative biodiversity monitoring approach to MAAT surveys that reduces reliance on taxonomic experts, requires less time and equipment, and could broaden the temporal and spatial scope at which we conduct biodiversity surveys.

Our results demonstrate that eDNA can provide a low-impact alternative to MAAT surveys for detecting core red algal biodiversity and can fill in some gaps left by traditional survey methods. At both sites, the eDNA survey made fewer detections than the MAAT survey; however, composition results differed between survey types. These findings are consistent with other studies (Takahara et al. 2020, Wikston 2021), suggesting the efficacy of certain species detections may depend on the method used, and neither method alone is likely a comprehensive representation of the biodiversity present. The species uniquely detected using the eDNA survey had all been detected in previous years at the study sites (Saunders, unpublished data), thus we can reasonably assume these are true detections and not false positives. Equally, our eDNA survey failed to detect some species that the MAAT survey detected and thus cannot be considered as a gold standard as per our result. For a biodiversity survey to be considered exhaustive, a holistic approach using multiple methods may be required. To improve species detection, eDNA methods must be improved and standardized. This would include standardized sample collection, DNA extraction method, genetic markers and PCR protocols, and sequencing depth. Our results contribute to the expanding research stipulating that eDNA surveys will be a critical part of the movement towards more frequent and comprehensive biodiversity monitoring (Deiner et al. 2021, Lacoursiere-Roussel et al. 2016).

Cryptic, Microscopic, and Introduced Species Detection

An important benefit to using an eDNA molecular approach to survey red algae is its ability to detect species that visual surveys often overlook. One example of this is *Acrochaetium secundatum* (Lyngbye) Nägeli, a small, filamentous red alga typically <5 mm (Dawes & Mathieson 2008, Nageli & Cramer 1858) and a common epiphyte or epizoite in the NW Atlantic. We found *Acrochaetium secundatum* using eDNA at the Wallace Cove site, which was not detected with the MAAT survey. This species requires microscopic identification and can be easily overlooked or misidentified as is true for other species of *Acrochaetium*.

Another unique detection by eDNA includes *Antithamnionella floccosa* (O.F.Müller) Whittick, which was found with eDNA at both sites but was only detected with the MAAT survey at the Lepreau site. This species is a low intertidal to shallow subtidal species and may be difficult to detect using intertidal visual surveys unless there is an exceptionally low tide, or with SCUBA. *Phymatolithon* sp.3AT crust and *Clathromorphum circumscriptum* (Strömfelt) Foslie /sp.1 circumscriptum were detected at our Lepreau site using eDNA but were not detected in our MAAT survey. Coralline crusts can be difficult to identify in the field and can look sufficiently similar to each other to be overlooked or discounted as a single entity in traditional floristic surveys. The addition of crustose corallines data to our SSU reference library was a particular advantage of using a rock scrape eDNA method as opposed to another eDNA technique.

Importantly, we detected *Bonnemaisonia hamifera* using eDNA at Wallace Cove, which was not detected in the MAAT survey. *Bonnemaisonia hamifera* is an invasive species whose global expansion has been rapid and aggressive, growing nearly

continuously over other native macroalgae (Garbary et al. 2020, Sadogurskiy et al. 2023). This species should be monitored for range expansion, and according to our results it may require an eDNA survey to do so.

Our eDNA survey detected *Coccotylus brodiei* (Turner) Kutzing at the Lepreau site, which was not detected with our MAAT approach. *Coccotylus brodiei* has highly variable morphology and can be confused with both *Coccotylus truncatus* (Pallas) M. J. Wynne & J.N. Heine and *Phyllophora pseudoceranoides* (S.G. Gremlin) Newroth & A.R.A. Taylor ex P.S. Dixon & L, M. Irvine in the field (Zgrundo & Zloch 2022). Both of these species have overlapping range and habitat with *C. brodiei.* The ability to detect these similar species is a distinct advantage over MAAT surveys.

Non-detections and Possible False Negatives

It should be noted explicitly that we used detections and non-detections in this survey, but that criterion is not equivalent to presence/absence. A non-detection in our eDNA survey does not mean that a taxon is absent from our sites, but it went undetected in our survey due to lack of reference material, insufficient sampling, or true absence from the site or extracted material.

There are some potential explanations for the detection or non-detection of certain species using eDNA and MAAT survey types based on the sampling approach and unique limitations of each method. Some species that we detected using eDNA were not found with the MAAT survey, possibly because of their cryptic nature or difficulty to find and identify. However, there are other unique detections from these surveys that are more difficult to explain. For example, we detected *Neopyropia leucosticta* (Thuret) L.E.

Yang & J. Brodie using eDNA at the Wallace Cove site but were unable to detect this species at the Lepreau site, despite having a reference sequence and having confirmed its presence with a same-season MAAT survey. *Neopyropia* can be found throughout the intertidal zone to the shallow subtidal (Kim et al. 2022), and thus is distributed within our sampling area. This non-detection likely represents an actual absence in our extracted material. Similarly, *Euthora cristata* (C.Agardh) J.Agardh and *Ptilota serrata* Kützing went undetected in our eDNA surveys but were found with both MAAT surveys. These species are largely subtidal and may have been present at the site but absent in our extracted material due to the intertidal limitations of our sampling approach. It is possible that we could detect these species with a more targeted sampling approach, or increased sampling, given that we detected other low intertidal to subtidal species with the eDNA approach (e.g., *Palmaria palmata* (Linnaeus) F.Weber & D.Mohr).

Species Resolving Power in the SSU

In our eDNA results we used both ecological and data context to make decisions surrounding some ambiguous identifications. For example, we had a genus-level match to a *Peyssonnelia* in our eDNA survey, but the only known species of this genus in our flora is *P. rosenvingei* F.Schmitz. We did not have a reference sequence for *P. rosenvingei but* using ecological context we concluded the genus-level *Peyssonnelia* sequence was our local species, *P. rosenvingei.*

Other eDNA identifications required further investigation, revealing the limitations of the species resolving power of the SSU. The SSU reference sequences for *Clathromorphum circumscriptum* and *Clathromorphum* sp.1 circumscriptum are

identical, however the two isolates differ in COI-5P, *psb*A, and *rbc*L (Saunders, unpublished data) and are treated as separate genetic groups in our database and in BOLD. Because the SSU sequences are impossible to distinguish, we detected at least one of the two *Clathromorphum* genetic groups in our eDNA survey, but we cannot confirm that both were present. To avoid potentially artificially inflating the differences between the two surveys, we combined this sequence to represent *Clathromorphum circumscriptum/Clathromorphum* sp. 1circumscriptum and we count this as a single detection in our eDNA results (Table 1.3). Our MAAT surveys only detected one of these genetic groups, *Clathromorphum circumscriptum*, suggesting our eDNA detection is also likely *Clathromorphum circumscriptum*, but another marker is required to confirm if both genetic groups are present in our eDNA survey.

We had an identical match to a *Polysiphonia stricta* sequence from Europe, likely *Polysiphonia* sp. 2stricta, which is also common in the Canadian Atlantic regions. Thus, we concluded this match to be *Polysiphonia* sp. 2stricta, despite not having a direct reference sequence for this species. Similarly, we do not have a reference sequence for *Polysiphonia* sp. 1stricta and *Polysiphonia* sp. 3stricta, however we have an eDNA sequence that matches *Polysiphonia* sp. Upon further inspection, this SSU sequence is slightly different in the variable regions to the sequence for *Polysiphonia* sp. 2stricta, and thus is likely either *Polysiphonia* sp. 1stricta or *Polysiphonia* sp. 3stricta. All three genetic groups of *Polysiphonia* are common in the Bay of Fundy (Saunders, unpublished data), and thus we concluded that our eDNA survey detected *Polysiphonia* sp. 2stricta, and either *Polysiphonia* sp. 1stricta or *Polysiphonia* sp. 3stricta, potentially both. In our eDNA results, we counted this as two *Polysiphonia* detections, one for *Polysiphonia* sp.

2stricta and one for *Polyisphonia* sp. 1stricta/*Polysiphonia* sp. 3stricta. Our Wallace Cove MAAT survey detected both *Polysiphonia* sp. 1stricta and *Polysiphonia* sp. 3stricta, but not *Polysiphonia* sp. 2stricta. Our Lepreau MAAT survey detected *Polysiphonia* sp. 2stricta and *Polysiphonia* sp. 1stricta. This confirms that all three *Polysiphonia* species were present at our sites, but improved reference material is required to distinguish between *Polysiphonia* sp. 1stricta and *Polysiphonia* sp. 3stricta in an eDNA survey.

Another identification of interest was the genus *Phycodrys.* We had an eDNA sequence match to *Phycodrys rubens* at our Lepreau site, where *P. rubens* has not been confirmed by MAAT survey to inhabit this region of the Bay of Fundy. This match is likely *Phycodrys* sp. 1NB, which is common at this site.

An eDNA sequence matched to *Corallina* sp. and is different from our reference *Corallina officinalis* sequences. *Corallina officinalis* is the only known species of this genus in our flora, therefore this sequence may represent a new or introduced species of *Corallina* in our flora, or it may be a contaminant or a result of sequencing error. The sequence appeared independently in our results four times, which surpassed our defined frequency thresholds; however, if we were to increase our threshold slightly, this sequence would have been discarded. Because the sequence is different from our other *Corallina* sequences, we included *Corallina* sp. in our identifications.

Limitations

Marker Selection

The accuracy and effectiveness of DNA metabarcoding technology depends greatly on the genetic marker and primer sets selected for the study (Chenuil 2006, Zhan

et al. 2014). Marker selection requires the consideration of inter- and intraspecific variation in the gene (species resolving power), quantity and quality of barcodes in the reference library, reputable primer sets, and amplification ability (Yang et al. 2016, Zhan et al. 2014).

There is currently no universal barcoding region for red algae (Saunders 2005); multiple markers and primer sets have been trialed, making comparison across studies and the development of large reference databases difficult and slow-going. Bates et al. (2007) found that identification to genus level resulted in only a 5% reduction in taxonomic information compared to high-resolution data. This suggests strict species-level resolution may not always be required for algal biodiversity surveys. This offers a potential way to overcome some of the difficulties related to developing species level reference libraries for the group of study, as only one reference sequence per genus would need to be generated.

With the previous in mind, we chose the SSU rRNA gene which satisfies the requirements for this study. The SSU is one of the most frequently used markers in phylogenetic studies across phyla (CD Genomics 2018, Khaw et al. 2020, Saunders & Moore 2013, Wang et al. 2017), and is therefore helpful for comparisons across taxonomic groups. Further, many reference sequences were already available in the PR^2 database and in GenBank, reducing time and expenses in developing a reference library.

Despite the utility of the SSU, some algal species differ in this gene by as little as a few nucleotides (Saunders & Kraft 1994), which can make differentiating between sequencing errors and true species difficult or impossible. Rare species and low abundance sequences should be carefully examined, and ideally confirmed with an

additional molecular marker. Additionally, the full-length SSU primers amplify all eukaryotes (Hendriks et al. 1989, Medlin et al. 1988), and therefore a percentage (32–45% per sample) of the raw reads were not red algal. This was an expected result given our marker selection and sampling approach.

Curated Reference Libraries

In addition to marker selection, the usefulness and value of eDNA metabarcoding is also limited by the quality and accessibility of high-quality reference databases (Coissac et al. 2016, Sacco et al. 2022). For optimal resolution, the database should contain sequences for the target region from multiple individuals of the same species with associated metadata and species information. These databases are a work in progress for all groups, especially within the marine algae where a standard barcoding region has yet to be decided (Saunders 2005) and nomenclature is constantly being updated with molecular data (Bolton 2020).

There are a few options when choosing a reference database for a metabarcoding study: 1) public databases, 2) curated or taxa-specific databases, and 3) supplemented databases. Public databases, such as GenBank and the Barcode of Life Database (BOLD), are free and accessible, and they span a large breadth of genetic markers and taxa. Using a public database eliminates the need to generate reference sequences, which is cost- and time-effective. However, public databases come with a series of inherent risks as some of the entries may not be up to date with new taxonomic information. For example, *Ceramium virgatum* Roth was identified using a public database as *Ceramium rubrum* C. Agardh, a name that has since been synonymized with *Ceramium virgatum* (Maggs et al. 2002). Additionally, reference sequences should be verified by a qualified taxonomist to

confirm the identities of vouchered material, which is often not the case in public databases. In 2014, less than 10% of the barcodes uploaded to BOLD were associated with formally described species (Taylor & Harris 2012), and many barcodes uploaded have no associated vouchered material for reference. This reinforces the need for expert taxonomists in every field to generate and verify the barcodes being used for metabarcoding studies further reenforcing that metabarcoding is a complementary method to traditional surveys, not a complete replacement.

Taxa-specific databases only include sequences for any taxa that are expected in the study system and exclude as many others as possible to minimize highly similar sequences in the search space (Dreier et al. 2019). These databases can significantly reduce computation time; however, they do not account for unexpected species, including those that may be newly introduced or have expanded outside their usual range. Taxa-specific databases tend to perform best in specific and restricted ecosystems or environments, e.g., the human mouth (Chen et al. 2010). For this study, we supplemented the PR^2 public database with regionally specific red algal barcodes. The identities of the sequences we added have been confirmed by either another molecular marker (*rbc*L or COI-5P) and/or morphological identification, thereby enhancing the reliability of the identifications made with the PR^2 RED1 library than if we had used strictly the publicly available database. This decision allowed us to capitalize on existing barcodes while filling in some of the regional gaps with reliable reference sequences. This supplemented database allows for the detection of species we would not necessarily expect to find in the Bay of Fundy, in the case of significant range expansion or a newly introduced species.

Bioinformatics

Despite the increasing interest and popularity of long-read technologies (Amarasinghe et al. 2020), most bioinformatic tools are designed and developed for short-read platforms like Illumina MiSeq (Sahraei et al. 2022). In 2020, the only tool to manage the entire pipeline of PacBio Sequel raw sequence data was SMRT Link, an unadaptable, computationally intensive, Linux-only distribution. The inaccessibility of the platform has sparked the development of many long-read tools and the adaptation of existing scripts for short-read amplicons to account for the different characteristics of PacBio data.

There is no universal bioinformatic pipeline, and no universal threshold at which species diverge (Sahraei et al. 2022), and thus choosing programs and parameters to run bioinformatic analyses relies on deep knowledge of the study system and marker of choice. Since our marker is a highly conserved gene, we wanted strict parameters for our query sequences to match the reference sequences. E-value (expectation value) is the number of expected hits that could be found by chance; it is dependent on the query sequence length and size of the database (NCBI, 2023). A smaller E-value represents a better match and can be used as an initial quality filter; with this in mind, we chose a low E-value threshold of 1×10^{-15}, and a high percent nucleotide sequence identity of 99.5%. Our query sequences were not manually edited in any way, meaning there could be some ambiguity at the ends of the sequences as sequencing errors frequently occur in the 3’ and 5’ ends (Jiao et al. 2013). These regions may introduce some noise; however, removal of all ambiguous regions can result in the deletion of valuable and genuine signal. We left these ambiguous regions untouched because of the strict parameters we had set in

downstream analyses. We also removed singletons and doubletons in our analyses to account for PCR error, removing any unique sequence that had an abundance of less than 3. A sequence that only appeared once or twice in thousands of reads is likely not representative of genuine diversity; however, there is always a risk of removing true diversity when discarding any low read sequences (Clare et al. 2016). There are no explicit rules when it comes to singleton removal; however, it is known that singleton removal reduces chimera content and computation time (Auer et al. 2017), and so at the risk of discarding some potentially rare species diversity, we chose to remove singletons and doubletons. Additionally, a study by Jiao et al. (2013) concluded that even though CCS output is meant to be error-corrected, approximately 20% of CCS reads are still considered low-quality in standard assembly programs and can still contain chimeras. Some protocols do not require intensive quality control steps on CCS outputs, but we still eliminated low read count sequences, and performed chimera removal to account for this 20% low quality.

Future Work

Future work must be done to continue developing high-quality reference libraries for red algae. As more sequences are generated, databases will become more reliable and useful, making the entire survey even quicker and more comprehensive. A benefit to the rock scrape method is that the dried material we used for extraction can be preserved and reanalyzed with a different marker or be used to monitor whole communities. Additionally, the raw sequence data can be rerun through an updated pipeline with a better library once these tools are further developed. Many red algal species complexes remain to be resolved,

and so taxonomic work must continue in this group to generate molecular data and associated metadata. Taxonomy is required to identify species and generate barcodes, so we can only identify via eDNA metabarcoding what we know via taxonomy. Generally, future studies in this field should work to standardize the sample collection and processing, genetic markers, and bioinformatic pipelines. Further, genetic markers and associated primers should be trialed to optimize species resolution and amplification ability. This study used SSU because of its accessibility (i.e., reference sequences available, primers available, amplification ability); however, due to the highly conserved nature of the SSU, other markers (e.g., *rbc*L, ITS, LSU) are likely to provide improved species resolution (Evans et al. 2007, Harper & Saunders 2001) within the Rhodophyta.

Standardization of DNA extraction methods, lab methods, and bioinformatic pipelines must be improved before eDNA metabarcoding can be considered a replacement tool for conducting biodiversity surveys. Future projects should strive towards developing standard methods throughout every step to achieve the most comprehensive results.

Conclusion

Marine ecosystem biodiversity continues to decline due to overharvesting, climate change, and other anthropological factors, making biodiversity monitoring an important tool for conservation. eDNA metabarcoding offers a hopeful alternative to current biodiversity monitoring protocols, which can be labor-intensive and reliant on a declining pool of taxonomic experts. Our results demonstrate that eDNA can detect core local red algal diversity (>50% of the red algal species detected using MAAT surveys) and can

detect cryptic red algal species that are hard to detect using traditional methods. Our eDNA survey did not outperform the MAAT survey in terms of species richness; however, it does provide a way to perform baseline biodiversity monitoring quickly and without taxonomic expertise. Allowing non-experts to conduct biodiversity monitoring could result in increased monitoring frequency in regions without local experts. This study is in accordance with other eDNA studies in all fields of study, supporting the idea that eDNA metabarcoding can be used to supplement and expand current monitoring practices.

Bibliography

Abdelgawad, A. M., El-Naggar, M. E., Elsherbiny, D. A., Ali, S., Abdel-Aziz, M. S., & Abdel-Monem, Y. K. (2020). Antibacterial carrageenan/cellulose nanocrystal system loaded with silver nanoparticles, prepared via solid-state technique. *Journal of Environmental Chemical Engineering*, *8*(5), 104276. https://doi.org/10.1016/j.jece.2020.104276

Agnarsson, I., & Kuntner, M. (2007). Taxonomy in a Changing World: Seeking Solutions for a Science in Crisis. *Systematic Biology*, *56*(3), 531–539. https://doi.org/10.1080/10635150701424546

Amarasinghe, S. L., Su, S., Dong, X., Zappia, L., Ritchie, M. E., & Gouil, Q. (2020). Opportunities and challenges in long-read sequencing data analysis. *Genome Biology*, *21*(1), 30. https://doi.org/10.1186/s13059-020-1935-5

Aragona, M., Haegi, A., Valente, M. T., Riccioni, L., Orzali, L., Vitale, S., Luongo, L., & Infantino, A. (2022). New-Generation Sequencing Technology in Diagnosis of Fungal Plant Pathogens: A Dream Comes True? *Journal of Fungi*, *8*(7), Article 7. https://doi.org/10.3390/jof8070737

Archer, A. W. (2013). World's highest tides: Hypertidal coastal systems in North America, South America and Europe. *Sedimentary Geology 284*(1), 1-25

Auer, L., Mariadassou, M., O'Donohue, M., Klopp, C., & Hernandez-Raquet, G. (2017). Analysis of large 16S rRNA Illumina data sets: Impact of singleton read filtering on microbial community description. *Molecular Ecology Resources*, *17*(6), e122–e132. https://doi.org/10.1111/1755-0998.12700

Baker, C. S., Steel, D., Nieukirk, S., & Klinck, H. (2018). Environmental DNA (eDNA) from the wake of the whales: droplet digital PCR for detection and species identification. *Frontiers in Marine Science*, 133.

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

Barnes, M. A., Turner, C. R., Lerde, C. L., Renshaw, M. A., Chadderton L. W. & Lodge, D. M. Environmental Conditions Influence eDNA Persistence in Aquatic Systems. (2014). *Environmental Science & Technology*. 2014;48(3):1819-27. https://pubs.acs.org/doi/10.1021/es404734p

Bates, C. R., Chopin, T., & Saunders, G. W. (2001). Monitoring seaweed diversity in the Bay of Fundy, New Brunswick, Canada. In *Opportunities and challenges for protecting, restoring, and enhancing coastal habitats in the Bay of Fundy. Proceedings of the 4th Bay of Fundy Science Workshop* (pp. 19-21).

Bates, C. R., Scott, G., Tobin, M., & Thompson, R. (2007). Weighing the costs and benefits of reduced sampling resolution in biomonitoring studies: Perspectives from the temperate rocky intertidal. *Biological Conservation, 137*(4), 617–625. https://doi.org/ 10.1016/j.biocon.2007.03.019

Bax, N., Williamson, A., Aguero, M., Gonzalez, E., & Geeves, W. (2003). Marine invasive alien species: A threat to global biodiversity. *Marine Policy, 27*(4), 313–323.

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2013). GenBank. *Nucleic Acids Research*, *41*(Database issue), D36–D42. https://doi.org/10.1093/nar/gks1195

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., Yu, D. W., & de Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution*, *29*(6), 358–367. https://doi.org/10.1016/j.tree.2014.04.003

Bolton, J. J. (2020). The problem of naming commercial seaweeds. *Journal of Applied Phycology*, *32*(2), 751–758. https://doi.org/10.1007/s10811-019-01928-0

Camacho, C., Coulouris, G. & Avagyan, V. (2009). BLAST+: architecture and applications. BMC Bioinformatics 10, 421. https://doi.org/10.1186/1471-2105-10-421

CD Genomics. (2018). 18S rRNA and Its Use in Fungal Diversity Analysis | CD Genomics Blog. Retrieved January 19, 2023, from https://www.cd-genomics.com/blog/18s-rrna-and-its-use-in-fungal-diversity-analysis/

Chankaew, W., Sakset, A., Chankaew, S., & Kumano, S. (n.d.). *Possibility of freshwater red algae as bioindicator for water quality of streams in Nakhon Si Thammarat province, Thailand*.

Chen, T., Yu, W.-H., Izard, J., Baranova, O. V., Lakshmanan, A., & Dewhirst, F. E. (2010). The Human Oral Microbiome Database: A web accessible resource for investigating oral microbe taxonomic and genomic information. *Database: The Journal of Biological Databases and Curation*, *2010*, baq013. https://doi.org/10.1093/database/baq013

Chenuil, A. (2006). Choosing the right molecular genetic markers for studying biodiversity: From molecular evolution to practical aspects. *Genetica*, *127*(1–3), 101–120. https://doi.org/10.1007/s10709-005-2485-1

Chung, N. C., Miasojedow, B., Startek, M., & Gambin, A. (2018) Jaccard: Test similarity between binary data using Jaccard/Tanimoto coefficients. https://cran.r-project.org/web/packages/jaccard/index.html

Clare, E. L., Chain, F. J., Littlefair, J. E., & Cristescu, M. E. (2016). The effects of parameter choice on defining molecular operational taxonomic units and resulting ecological analyses of metabarcoding data. *Genome*, *59*(11), 981-990.

Coissac, E., Hollingsworth, P. M., Lavergne, S., & Taberlet, P. (2016). From barcodes to genomes: Extending the concept of DNA barcoding. *Molecular Ecology*, *25*(7), 1423–1428. https://doi.org/10.1111/mec.13549

Collins, R. A., Wangensteen, O. S., O'Gorman, E. J., Mariani, S., Sims, D. W., & Genner, M. J. (2018). Persistence of environmental DNA in marine systems. *Communications Biology*, *1*(1), Article 1. https://doi.org/10.1038/s42003-018-0192-6

Craig, R. K. (2012). Marine Biodiversity, Climate Change, and Governance of the Oceans. *Diversity*, *4*(2), Article 2. https://doi.org/10.3390/d4020224

Dawes, C. J., & Mathieson, A. C. (2008). *The seaweeds of Florida*. University Press of Florida.

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D. M., de Vere, N., Pfrender, M. E., & 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

Desplanque, C., & Mossman, D. J. (2001). Bay of Fundy Tides. *Geoscience Canada*. https://journals.lib.unb.ca/index.php/GC/article/view/4065

Dreier, M., Meola, M., Berthoud, H., Shani, N., Wechsler, D., & Junier, P. (2022). High-throughput qPCR and 16S rRNA gene amplicon sequencing as complementary methods for the investigation of the cheese microbiota. *BMC microbiology*, *22*(1), 48.

Evans, K. M., Wortley, A. H., & Mann, D. G. (2007). An Assessment of Potential Diatom "Barcode" Genes (cox1, rbcL, 18S and ITS rDNA) and their Effectiveness in Determining Relationships in *Sellaphora* (Bacillariophyta). *Protist*, *158*(3), 349–364. https://doi.org/10.1016/j.protis.2007.04.001

FAO, C. on G. R. for F. and A. (2019). *The state of the world's biodiversity for food and agriculture*. http://www.fao.org/state-of-biodiversity-for-food-agriculture/en/

Gagnon, K., McKindsey, C. W., & Johnson, L. E. (2015). Roles of dispersal mode, recipient environment and disturbance in the secondary spread of the invasive seaweed *Codium fragile. Biological Invasions*, *17*(4), 1123–1136. https://doi.org/10.1007/s10530-014-0782-2

Gamfeldt, L., Lefcheck, J. S., Byrnes, J. E. K., Cardinale, B. J., Duffy, J. E., & Griffin, J. N. (2015). Marine biodiversity and ecosystem functioning: What's known and what's next? *Oikos*, *124*(3), 252–265. https://doi.org/10.1111/oik.01549

Garbary, D. J., D'Archino, R., Flack, B., Hepburn, C. D., Nelson, W. A., Pritchard, D., & Sutherland, J. E. (2020). First record of *Bonnemaisonia hamifera* (Bonnemaisoniales,

Rhodophyta) in the South Pacific, from the South Island of New Zealand. *New Zealand Journal of Marine and Freshwater Research, 54*(2), 167-176.

Gold, Z., Sprague, J., Kushner, D. J., Marin, E. Z., & Barber, P. H. (2021). EDNA metabarcoding as a biomonitoring tool for marine protected areas. *PLOS ONE, 16*(2), e0238557. https://doi.org/10.1371/journal.pone.0238557

Govindarajan, A. F., McCartin, L., Adams, A., Allan, E., Belani, A., Francolini, R., Fujii, J., Gomez-Ibañez, D., Kukulya, A., Marin, F., Tradd, K., Yoerger, D. R., McDermott, J. M., & Herrera, S. (2022). Improved biodiversity detection using a large-volume environmental DNA sampler with in situ filtration and implications for marine eDNA sampling strategies. *Deep Sea Research Part I: Oceanographic Research Papers, 189*, 103871. https://doi.org/10.1016/j.dsr.2022.103871

Government of New Brunswick, C. (2012). *Bay of Fundy Marine Aquaculture Site Allocation Policy*. https://www2.gnb.ca/content/gnb/en/departments/10/aquaculture/content/site_-cation_policy.html#8

Grealy, A. C., McDowell, M. C., Scofield, P., Murray, D. C., Fusco, D. A., Haile, J., Prideaux, G. J., & Bunce, M. (2015). A critical evaluation of how ancient DNA bulk bone metabarcoding complements traditional morphological analysis of fossil assemblages. *Quaternary Science Reviews, 128*, 37–47. https://doi.org/10.1016/j.quascirev. 2015.09.014

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., de Vargas, C., Decelle, J., del Campo, J., Dolan, J. R., Dunthorn, M., Edvardsen, B., Holzmann, M., Kooistra, W. H. C. F., Lara, E., Le Bescot, N., Logares, R. &

Christen, R. (2013). The Protist Ribosomal Reference database (PR2): A catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Research*, *41*(Database issue), D597–D604. https://doi.org/10.1093/nar/gks1160

Harper, J. T., & Saunders, G. W. (2001). Molecular systematics of the Florideophyceae (Rhodophyta) using nuclear large and small subunity rDNA sequence data. *Journal of Phycology*, *37*(6), 1073–1082. https://doi.org/10.1046/j.1529-8817.2001.00160.x

Hendriks, L., Goris, A., Neefs, J.-M., Van De Peer, Y., Hennebert, G., & De Wachter, R. (1989). The Nucleotide Sequence of the Small Ribosomal Subunit RNA of the Yeast *Candida albicans* and the Evolutionary Position of the Fungi among the Eukaryotes. *Systematic and Applied Microbiology*, *12*(3), 223–229. https://doi.org/10.1016/S0723-2020(89)80066-9

Hon, T., Mars, K., Young, G., Tsai, Y.-C., Karalius, J. W., Landolin, J. M., Maurer, N., Kudrna, D., Hardigan, M. A., Steiner, C. C., Knapp, S. J., Ware, D., Shapiro, B., Peluso, P., & Rank, D. R. (2020a). Highly accurate long-read HiFi sequencing data for five complex genomes. *Scientific Data*, *7*(1), Article 1. https://doi.org/10.1038/s41597-020-00743-4

Hufnagel, D. E., Hufford, M. B., & Seetharam, A. S. (2020). SequelTools: A suite of tools for working with PacBio Sequel raw sequence data. *BMC Bioinformatics*, *21*(1), 429. https://doi.org/10.1186/s12859-020-03751-8

Jaccard, P. (1908). Nouvelles recherches sur la distribution florale. *Bull. Soc. Vaud. Sci. Nat.*, *44*, 223-270.

Jiao, X., Zheng, X., Ma, L., Kutty, G., Gogineni, E., Sun, Q., Sherman, B. T., Hu, X., Jones, K., Raley, C., Tran, B., Munroe, D. J., Stephens, R., Liang, D., Imamichi, T., Kovacs, J. A., Lempicki, R. A., & Huang, D. W. (2013). A Benchmark Study on Error Assessment and Quality Control of CCS Reads Derived from the PacBio RS. *Journal of Data Mining in Genomics & Proteomics*, *4*(3), 16008. https://doi.org/10.4172/2153-0602.1000136

Jo, H., Ventura, M., Vidal, N., Gim, J., Buchaca, T., Barmuta, L. A., Jeppesen, E., & Joo, G. (2015). Discovering hidden biodiversity: The use of complementary monitoring of fish diet based on DNA barcoding in freshwater ecosystems. *Ecology and Evolution*, *6*(1), 219–232. https://doi.org/10.1002/ece3.1825

Kalu, E. I., Reyes-Prieto, A., & Barbeau, M. A. (2023). Community dynamics of microbial eukaryotes in intertidal mudflats in the hypertidal Bay of Fundy. *ISME Communications*, *3*(1), Article 1. https://doi.org/10.1038/s43705-023-00226-8

Khaw, Y. S., Khong, N. M. H., Shaharuddin, N. A., & Yusoff, F. M. (2020). A simple 18S rDNA approach for the identification of cultured eukaryotic microalgae with an emphasis on primers. *Journal of microbiological methods*, *172*, 105890.

Knudsen, S. W., Hesselsøe, M., Thaulow, J., Agersnap, S., Hansen, B. K., Jacobsen, M. W., Bekkevold, D., Jensen, S. K. S., Møller, P. R., & Andersen, J. H. (2022). Monitoring of environmental DNA from nonindigenous species of algae, dinoflagellates and animals in the North East Atlantic. *Science of The Total Environment*, *821*, 153093. https://doi.org/10.1016/j.scitotenv.2022.153093

Kobayashi, M., & Fujita, D. (2014). Can thallus color of red algae be used as an environmental indicator in shallow waters? *Journal of applied phycology*, *26*, 1123-1131.

Lacoursière-Roussel, A., Howland, K., Normandeau, E., Grey, E. K., Archambault, P., Deiner, K., Lodge, D. M., Hernandez, C., Leduc, N., & Bernatchez, L. (2018). EDNA metabarcoding as a new surveillance approach for coastal Arctic biodiversity. *Ecology and Evolution*, *8*(16), 7763–7777. https://doi.org/10.1002/ece3.4213

Le Gall, L., & Saunders, G. W. (2010). Dna barcoding is a powerful tool to uncover algal diversity: A case study of the Phyllophoraceae (Gigartinales, Rhodophyta) in the Canadian flora 1. *Journal of phycology*, *46*(2), 374-389.

Leray, M., Knowlton, N., Ho, S. L., Nguyen, B. N., & Machida, R. J. (2019). GenBank is a reliable resource for 21st century biodiversity research. *Proceedings of the National Academy of Sciences*, *116*(45), 22651-22656.

Li, W., Jaroszewski, L., & Godzik, A. (2001). Clustering of highly homologous sequences to reduce the size of large protein databases. *Bioinformatics (Oxford, England)*, *17*(3), 282–283. https://doi.org/10.1093/bioinformatics/17.3.282

Limin Fu, Beifang Niu, Zhengwei Zhu, Sitao Wu & Weizhong Li. (n.d.-a). *CD-HIT: accelerated for clustering the next generation sequencing data. Bioinformatics, (2012), 28 (23): 3150-3152. Doi: 10.1093/bioinformatics/bts565. PDF, pubmed.*

Lindenmayer, D. B., Gibbons, P., Bourke, M., Burgman, M., Dickman, C. R., Ferrier, S., Fitzsimons, J., Freudenberger, D., Garnett, S. T., Groves, C., Hobbs, R. J., Kingsford, R. T., Krebs, C., Legge, S., Lowe, A. J., Mclean, R., Montambault, J., Possingham, H., Radford, J. & Zerger, A. (2012). Improving biodiversity

monitoring. *Austral Ecology*, *37*(3), 285–294. https://doi.org/10.1111/j.1442-9993.2011.02314.x

Loreau, M., Naeem, S., Inchausti, P., Bengtsson, J., Grime, J. P., Hector, A., Hooper, D. U., Huston, M. A., Raffaelli, D., Schmid, B., Tilman, D., & Wardle, D. A. (2001). Biodiversity and Ecosystem Functioning: Current Knowledge and Future Challenges. *Science*, *294*(5543), 804–808.

Maggs, C. A., Ward, B. A., McIvor, L. M., Evans, C. M., Rueness, J., & Stanhope, M. J. (2002). Molecular analyses elucidate the taxonomy of fully corticated, nonspiny species of *Ceramium* (Ceramiaceae, Rhodophyta) in the British Isles. *Phycologia*, *41*(4), 409–420. https://doi.org/10.2216/i0031-8884-41-4-409.1

Mas-Carrió, E., Schneider, J., Nasanbat, B., Ravchig, S., Buxton, M., Nyamukondiwa, C., Stoffel, C., Augugliaro, C., Ceacero, F., Taberlet, P., Glaizot, O., Christe, P., & Fumagalli, L. (2022). Assessing environmental DNA metabarcoding and camera trap surveys as complementary tools for biomonitoring of remote desert water bodies. *Environmental DNA*, *4*(3), 580–595. https://doi.org/10.1002/edn3.274

Mathieson, A. C., Dawes, C. J., Pederson, J., Gladych, R. A., & Carlton, J. T. (2008). The Asian red seaweed *Grateloupia turuturu* (Rhodophyta) invades the Gulf of Maine. *Biological Invasions*, *10*(7), 985–988. https://doi.org/10.1007/s10530-007-9176-z

Medlin, L., Elwood, H. J., Stickel, S., & Sogin, M. L. (1988). The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. *Gene*, *71*(2), 491–499. https://doi.org/10.1016/0378-1119(88)90066-2

Meiklejohn, K. A., Damaso, N., & Robertson, J. M. (2019). Assessment of BOLD and GenBank – Their accuracy and reliability for the identification of biological

materials. *PLOS ONE*, *14*(6), e0217084. https://doi.org/10.1371/journal.pone.0217084

Milla, L., Schmidt-Lebuhn, A., Bovill, J., & Encinas-Viso, F. (2022). Monitoring of honey bee floral resources with pollen DNA metabarcoding as a complementary tool to vegetation surveys. *Ecological Solutions and Evidence*, *3*(1), e12120. https://doi.org/10.1002/2688-8319.12120

Mineur, F., Arenas, F., Assis, J., Davies, A. J., Engelen, A. H., Fernandes, F., & De Clerck, O. (2015). European seaweeds under pressure: Consequences for communities and ecosystem functioning. *Journal of sea research*, *98*, 91-108.

Minhas, A., Kaur, B., & Kaur, J. (2020). Genomics of algae: Its challenges and applications. *Pan-genomics: Applications, Challenges, and Future Prospects*, 261-283.

Mora, C., & Sale, P. (2011). Ongoing global biodiversity loss and the need to move beyond protected areas: A review of the technical and practical shortcomings of protected areas on land and sea. *Marine Ecology Progress Series*, *434*, 251–266. https://doi.org/10.3354/meps09214

Moritz, C., & Cicero, C. (2004). DNA Barcoding: Promise and Pitfalls. *PLOS Biology*, *2*(10), e354. https://doi.org/10.1371/journal.pbio.0020354

Muthukrishnan, R., & Larkin, D. J. (2020). Invasive species and biotic homogenization in temperate aquatic plant communities. *Global Ecology and Biogeography*, *29*(4), 656–667. https://doi.org/10.1111/geb.13053

Nägeli, C. (1858). *Die Stärkekörner: Morphologische, physiologische, chemisch-physicalische und systematisch-botanische Monographie* (Vol. 2). Schulthess.

Nakagawa, H., Yamamoto, S., Sato, Y., Sado, T., Minamoto, T., & Miya, M. (2018). Comparing local- and regional-scale estimations of the diversity of stream fish using eDNA metabarcoding and conventional observation methods. *Freshwater Biology*, *63*(6), 569–580. https://doi.org/10.1111/fwb.13094

National Center for Biotechnology Information (NCBI)[Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; (1988) – [cited 2023 Jul 06]. Available from: https://www.ncbi.nlm.nih.gov/

Natural Resources Canada (2022). 2002-03 Estimates: A Report on Plans and Priorities. Retrieved from https://publications.gc.ca/site/eng/378085/publication.html.

Newton, C., Bracken, M. E. S., McConville, M., Rodrigue, K., & Thornber, C. S. (2013). Invasion of the Red Seaweed *Heterosiphonia japonica* Spans Biogeographic Provinces in the Western North Atlantic Ocean. *PLOS ONE, 8*(4), e62261 https://doi.org/1 0.1371/journal.pone.0062261

Niemelä, J. (2000). Biodiversity monitoring for decision-making. *Annales Zoologici Fennici*, *37*(4), 307–317.

Norton, T. A., Melkonian, M., & Andersen, R. A. (1996). Algal biodiversity. Phycologia, 35(4), 308-326.

Nylund, G. M., Enge, S., & Pavia, H. (2013). Costs and benefits of chemical defence in the red alga *Bonnemaisonia hamifera*. *PloS one*, *8*(4), e61291.

Palumbi, S. R., Sandifer, P. A., Allan, J. D., Beck, M. W., Fautin, D. G., Fogarty, M. J., Halpern, B. S., Incze, L. S., Leong, J.-A., Norse, E., Stachowicz, J. J., & Wall, D. H. (2009). Managing for ocean biodiversity to sustain marine ecosystem services.

Frontiers in Ecology and the Environment, *7*(4), 204–211. https://doi.org/10.1890/070135

Pikitch, E. K. (2018). A tool for finding rare marine species. *Science*, *360*(6394), 1180–1182. https://doi.org/10.1126/science.aao3787

Port, J. A., O'Donnell, J. L., Romero-Maraccini, O. C., Leary, P. R., Litvin, S. Y., Nickols, K. J., Yamahara, K. M., & 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

Prasanthi, N., Prasannakumar, C., Annadurai, D., & Mahendran, S. (2020). Identifying seaweeds species of Chlorophyta, Phaeophyta and Rhodophyta using DNA barcodes. *bioRxiv*, 2020-08.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *41*(Database issue), D590–D596. https://doi.org/10.1093/nar/gks1219

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

Ratnasingham, S., & Hebert, P. D. N. (2007). bold: The Barcode of Life Data System (http://www.barcodinglife.org). *Molecular Ecology Notes*, *7*(3), 355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x

Rees, H. C., Maddison, B. C., Middleditch, D. J., Patmore, J. R., & Gough, K. C. (2014). The detection of aquatic animal species using environmental DNA–a review of eDNA as a survey tool in ecology. *Journal of applied ecology*, *51*(5), 1450-1459.

Rinne, H., & Kostamo, K. (2022). Distribution and species composition of red algal communities in the northern Baltic Sea. *Estuarine, Coastal and Shelf Science, 269*, 107806. https://doi.org/10.1016/j.ecss.2022.107806

Rognes T, Flouri T, Nichols B, Quince C, Mahé F. (2016). (n.d.). *VSEARCH: a versatile open source tool for metagenomics. PeerJ 4:e2584. Doi: 10.7717/peerj.2584*.

Rourke, M. L., Fowler, A. M., Hughes, J. M., Broadhurst, M. K., DiBattista, J. D., Fielder, S., Walburn, J. W., Furlan E. M. (2022). Environmental DNA (eDNA) as a tool for assessing fish biomass: A review of approaches and future considerations for resource surveys. *Environmental DNA*, *4*(1), 9-33. https://onlinelibrary.wiley.com/doi/full/10.1002/edn3. 185

Saccò, M., Guzik, M. T., van der Heyde, M., Nevill, P., Cooper, S. J., Austin, A. D., ... & White, N. E. (2022). eDNA in subterranean ecosystems: Applications, technical aspects, and prospects. *Science of the Total Environment*, 153223.

Sadogurskiy, S. Y., Belich, T. V., & Sadogurskaya, S. A. (2023). Invasion of the alien species *Bonnemaisonia hamifera* Hariot in coastal Phytocenoses near the southern coast of Crimea (the Black Sea). *Inland Water Biology*, *16*(1), 81-86.

Sahraei, E. S., Furneaux, B., Kluting, K., Zakieh, M., Rydin, H., Hytteborn, H., & Rosling, A. (2022). Effects of operational taxonomic unit inference methods on soil microeukaryote community analysis using long read metabarcoding. *Ecology and Evolution*, *12*(3), e8676.

Saunders, G. W. (2005). Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. *Philosophical transactions of the Royal Society B: Biological sciences*, *360*(1462), 1879-1888.

Saunders, G. W. (2008). A DNA barcode examination of the red algal family Dumontiaceae in Canadian waters reveals substantial cryptic species diversity. 1. The foliose *Dilsea–Neodilsea* complex and *Weeksia*. *Botany*, *86*(7), 773–789. https://doi.org/10.1139/B08 001

Saunders, G. W., & McDevit, D. C. (2012). Methods for DNA Barcoding Photosynthetic Protists Emphasizing the Macroalgae and Diatoms. In W. J. Kress & D. L. Erickson (Eds.), *DNA Barcodes: Methods and Protocols* (pp. 207–222). Humana Press. https://doi.org/10.1007/978-1-61779-591-6_10

Saunders, G. W., & Kraft, G. T. (1994). Small-subunit rRNA gene sequences from representatives of selected families of the Gigartinales and Rhodymeniales (Rhodophyta). 1. Evidence for the Plocamiales ord. nov. *Canadian journal of botany*, *72*(9), 1250-1263.

Saunders, G. W., & Moore, T. E. (2013). Refinements for the amplification and sequencing of red algal DNA barcode and RedToL phylogenetic markers: a summary of current primers, profiles and strategies. *Algae*, *28*(1), 31-43.

Saunders, G. W., & Withall, R. D. (2006). Collections of the invasive species *Grateloupia turuturu* (Halymeniales, Rhodophyta) from Tasmania, Australia. *Phycologia*, *45*(6), 711-714.

Sayers, E. W., Cavanaugh, M., Clark, K., Ostell, J., Pruitt, K. D., & Karsch-Mizrachi, I. (2020). GenBank. Nucleic Acids Research, 48(D1), D84–D86.https://doi.org/ 10.1093/nar/gkz956

Statistics Canada (2021). Census Profile 2021. Retrieved from https://www12.statcan.gc.ca/ census-recensement/2021/dp-pd/prof/index.cfm?Lang=E

Storey JD, Bass AJ, Dabney A, Robinson D (2023). *qvalue: Q-value estimation for false discovery rate control*. R package version 2.32.0, http://github.com/jdstorey/qvalue

Szinte, A. L., Taylor, J. C., Abosede, A. T., & Vis, M. L. (2020). Current status of freshwater red algal diversity (Rhodophyta) of the African continent including description of new taxa (Batrachospermales). *Phycologia*, *59*(3), 187-199.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular ecology*, *21*(8), 2045-2050.

Takahara, T., Iwai, N., Yasumiba, K., & Igawa, T. (2020). Comparison of the detection of 3 endangered frog species by eDNA and acoustic surveys across 3 seasons. *Freshwater Science*, *39*(1), 18-27.

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*(3), 377-388.

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.

Vaulot, D. (2019). *pr2database/pr2database: PR2 version 4.12.0* (v4.12.0). Zenodo. https://doi.org/10.5281/ZENODO.3362765

Wang, C., Zhang, T., Wang, Y., Katz, L. A., Gao, F., & Song, W. (2017). Disentangling sources of variation in SSU rDNA sequences from single cell analyses of ciliates: Impact of copy number variation and experimental error. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1859), 20170425. https://doi.org/10.1098/rspb.2017.0425

Wang, X., He, L., Ma, Y., Huan, L., Wang, Y., Xia, B., & Wang, G. (2020). Economically important red algae resources along the Chinese coast: History, status, and prospects for their utilization. *Algal Research*, *46*, 101817. https://doi.org/10.1016/j.algal.2020.101817

Weizhong Li & Adam Godzik Bioinformatics, (2006) 22:1658-9. (n.d.). *Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences", Weizhong Li & Adam Godzik Bioinformatics, (2006) 22:1658-9*.

Wikston, M. J. (2021). *Comparative efficacy of eDNA and conventional methods for monitoring wetland anuran communities* (Doctoral dissertation, Trent University (Canada)).

Worm, B., Barbier, E. B., Beaumont, N., Duffy, J. E., Folke, C., Halpern, B. S., ... & Watson, R. (2006). Impacts of biodiversity loss on ocean ecosystem services. *science*, *314*(5800), 787-790.

Worm, B., & Lotze, H. K. (2021). Chapter 21—Marine biodiversity and climate change. In T. M. Letcher (Ed.), *Climate Change (Third Edition)* (pp. 445–464). Elsevier. https://doi.org/10.1016/B978-0-12-821575-3.00021-9

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*(1), 1-12.

Yang, E. C., Boo, S. M., Bhattacharya, D., Saunders, G. W., Knoll, A. H., Fredericq, S., ... & Yoon, H. S. (2016). Divergence time estimates and the evolution of major lineages in the florideophyte red algae. *Scientific reports*, *6*(1), 21361.

Yoon, H. S., Nelson, W., Lindstrom, S. C., Boo, S. M., Pueschel, C., Qiu, H., & Bhattacharya, D. (2017). Rhodophyta. In J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits (Eds.), *Handbook of the Protists* (pp. 89–133). Springer International Publishing. https://doi.org/10.1007/978-3-319-28149-0_33

Zgrundo, A., & Złoch, I. (2022). Gone and Back—The Anthropogenic History of *Coccotylus brodiei* (Turner) Kützing and *Furcellaria lumbricalis* (Hudson) JV Lamouroux in the Gulf of Gdańsk (Southern Baltic Sea). *Water*, *14*(14), 2181.

Zhan, A., Bailey, S. A., Heath, D. D., & Macisaac, H. J. (2014). Performance comparison of genetic markers for high-throughput sequencing-based biodiversity assessment in complex communities. *Molecular ecology resources*, *14*(5), 1049-1059.

Curriculum Vitae

Candidate's full name: Josephine Crichton

Universities attended: University of Alberta, Bachelor of Science with Specialization in Ecology, Evolution and Environmental Biology, 2020

Conference Presentations:

59th Northeast Algal Symposium, 2021. An Analysis of DNA Metabarcoding Techniques on Rhodophyta (Red Algae) Biodiversity and Phenology in the Bay of Fundy.

60th Northeast Algal Symposium, 2022. Monitoring Red Algal Biodiversity Using eDNA Metabarcoding.