UNIVERSITY
OF OSLO

Master thesis

The design and application of a qPCR assay for the study of light regime as a driver of rhodopsin expression in high Arctic phytoplankton.

Marine Molecular Biology

Emilie Nicholls

Master of Bioscience

Master thesis in Biology- Autumn 2023- BIOS5960 (60 ECTS)

University of Oslo

Faculty of Mathematics and Natural Sciences

Table of Contents

ABSTRACT **4**

INTRODUCTION **6**

1.0 Overview 6
2.0 Research location 7
3.0 Composition, utilization and biological effects of light 8
4.0 Rhodopsins 14

RESEARCH GOAL **17**

MATERIALS AND METHODS **19**

1.0 Sampling 19
2.0 Sample processing 21
3.0 Primer design 24
4.0 Primer testing 31
5.0 qPCR 36
6.0 PCR and qPCR product purification 39
7.0 Data Analysis 40

RESULTS **43**

Research Aim 1: Light regime (spectral composition, intensity, and duration). 43
Research Aim 2: Design of a qPCR assay through PCR optimization 45
Research aim 3: Drivers of rhodopsin expression 56

DISCUSSION **61**

Research Aim 1: Spectral composition throughout time series 61
Research Aim 2: Design of a qPCR assay through PCR optimization 62
Research Aim 3: Drivers of rhodopsin expression 68
Future Studies 74

CONCLUSION **75**

REFERENCES **76**

SUPPLEMENTARY MATERIAL **90**

Protocols 90
Physicochemical data 107
Primers 109
Table S6. Poly-A primers. 111
PCR trials 111
qPCR 115

Acknowledgments

I would like to thank my supervisors at UNIS Anna Vader and Ane Cecilie Kvernvik for granting me the opportunity to take on this Master project. Combining marine and molecular biology has long been a dream of mine, and being able to do so in a setting as spectacular as Svalbard is something I cherish deeply. I would also like to thank Stein Fredriksen who supported with great enthusiasm my move to Svalbard and offered his guidance and supervision throughout this thesis.

Special thanks to the lab technicians at UNIS, Stuart Thomson and Hana Spickova for both their assistance during field campaigns, and most importantly their words of reassurance and guidance in the lab when things did not go to plan.

I would not be where I am today if it weren't for the support from my parents and my older sister, so to them, Thank you, Merci, Obrigado.

The friends and office colleagues I have made in Oslo and at UNIS will forever remain a fundamental part of my master's degree. To the women with the force of nature, the women of science in Svalbard, Zoë, Saskia, Agnes, Annemijn, Charlotte, Julie, Anita, Catherine, and Elena, the support and guidance that you have brought me during this process cannot be underestimated. To the infamous 'Master of Disasters', your public disturbance kept me afloat throughout this thesis and I will fondly cherish the chaos amidst our most intensive writing periods.

To Mathias Morthen, who probably does not realize the gravity that his friendship, support, and guidance has contributed to the completion of this degree.

Finally, I would like to thank my body for pulling through despite the health ailments that I have experienced in the last 15 months whilst conducting this research. Receiving a diagnosis and major surgery for Stage 3 endometriosis in the middle of this thesis came with its fair share of difficulties and complications in the aftermath. Being able to complete this thesis despite being unable to stand on some days is a feat that I will always remember as a huge achievement.

Abstract

Several months of darkness during polar night impose unique challenges on polar phototrophic organisms, whereby solar energy capture via plastid-based photosystems remains unattainable for an extended period of time. Nevertheless, energy acquisition via the contribution of retinal-based phototrophy from microbial rhodopsins in high Arctic phytoplankton remains entirely unexplored. Light regime at the subsurface, via the investigation of spectral composition, intensity and duration, was studied throughout this time series spanning the transition from polar night to the spring equinox to explore this as a potential driver of rhodopsin expression. Blue- and green-light was observed as the dominant light type, especially during field campaigns associated with the lowest solar declination angles, suggesting that light-harvesting activity from blue- and green-light absorbing rhodopsins could be sustained given spectral composition and intensity.

A quantitative PCR assay based on the use of SYBR Green was developed to investigate rhodopsin gene abundance and expression in two strains of high Arctic phytoplankton as the light climate evolves. Preliminary testing via PCR was performed whilst continuously optimizing reaction conditions (primer concentrations, temperature trials, genetic material type and quantity). With the addition of a standard curve to standardize amplification, the assay was applied to DNA and RNA extracts from environmental samples collected bimonthly throughout the time series. Stochasticity and primer entropy was observed, likely due to their use on environmental samples containing very little genetic material. The assay did provide consistent results and demonstrated that stronger deductions can be made for rhodopsin gene abundance and expression in high Arctic phytoplankton should a higher yield of target DNA/RNA be present in the samples.

Ambient environmental conditions at the sampling site (temperature, salinity, nutrients, and fluorescence) displayed values expected for a prebloom phase and were investigated as drivers or limitations of rhodopsin expression. Sudden changes in subsurface water temperature and light intensity can be hypothesized as drivers which downregulate rhodopsin expression in the two high Arctic strains of phytoplankton which were investigated.

Keywords

Arctic, polar night, phytoplankton, rhodopsins, time series, qPCR assay.

Abbreviations

IsA	The Isfjorden Adventfjorden time series
WSC	West Spistsbergen Current
PAR	Photosynthetically Active Radiation
OTU	Operational Taxonomic Unit
PPR	Proton-pumping rhodopsin
HGT	Horizontal Gene Transfer
NCBI	National Center for Biotechnology Information
BLAST	Basic Local Alignment Search Tool
NEB	New England Biolabs®
Fp	Forward primer
Rp	Reverse primer
Chla	Chlorophyll *a*
Lhcf	Light-harvesting complex protein
psbA	D1 protein in PSII
rbcL	large subunit of ribulose-bisphosphate carboxylase
SC	Standard Curve
NTC	No Template Control

Introduction

1.0 Overview

The focus of this study are the understudied rhodopsin proteins present in eukaryotic phytoplankton. With functions ranging from phototaxis to flagellum rotation, they offer the potential of an alternative source of biogenic carbon formation in heterotrophs and phytoplankton. Rhodopsins occur widely in the surface ocean, having been found in 13 to 80% of marine bacteria and archaea (Béjà et al., 2000). However, compared to their bacterial counterparts, little knowledge has been gained of their abundance and expression patterns in eukaryotic phytoplankton, especially with regards to their requirements and biological use of light.

Until recently, it was assumed that polar waters were devoid of life and in a state of dormancy during polar night due to the lack of solar radiation. Concomitantly, the understanding of rhodopsin occurrence during this time is essentially non-existent. However, emerging research is redefining this period as rather characterized by a number of processes and interactions yet to be fully understood (Berge et al., 2015). Nevertheless, very few studies exist on the response of natural polar phytoplankton communities to light/ dark cycles, especially with regards to the use non-plastid-based photosystems. This study aims to be the first at 1. Designing a qPCR assay for rhodopsin in strains of high Arctic phytoplankton, and 2. Studying rhodopsin expression in said phytoplankton as the light regime at an Arctic location evolves from polar night to the spring equinox.

This introduction will first detail the study site- the Svalbard archipelago- before moving on to the complex seasonality of its light climate and inherent biological effects. Background concerning the fundamental aspects of light in photobiological studies will then be presented, followed by the phylogenetic origin of rhodopsin, and their prevalence in phytoplankton. Attention will be paid to the relevance of rhodopsin in polar environments through the mention of relevant studies whose findings associate rhodopsin occurrence with fitness-maximizing activities in the absence of light-based cues. Finally, the research aim of the work will be outlined.

2.0 Research location

Ecosystem dynamics

Svalbard, a Norwegian archipelago located roughly halfway between the northern coast of Norway and the North Pole (ranging from 74-81°N), is ecologically diverse and home to a plethora of food webs and ecosystem processes (Seuthe et al., 2011).
For the purpose of this study, I will focus on the ecosystem dynamics at m chosen site: the "IsA time series" sampling site at the Isfjorden/Adventfjorden interface (78°15'.669''N, 15°32'.023''E) (Fig. 1), which due to its proximity to the research facilities at the University Centre in Svalbard (UNIS), provides a model system for studying the Arctic and climate change. It hopes to build a comprehensive data set of a glacier/fjord/ current interface which may be utilized for future research in Arctic and climate change studies.

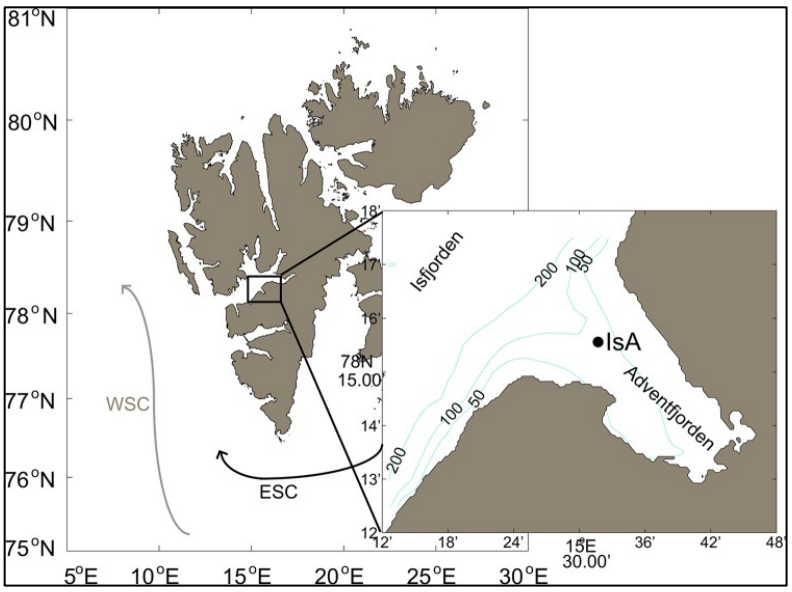


Figure 1. Map showing location of time series sampling site in Svalbard (Marquardt et al., 2016).

This sampling site was chosen as my field campaigns could be linked to the bimonthly sampling that occurs for the time series, in the hopes of contributing additional data of ecological relevance. Additionally, it enabled to take advantage of the existing information about community composition and seasonal ambient conditions at this site.

Isfjorden is the largest fjord on the West coast of Spitsbergen and is strongly influenced by the inflow of warm Atlantic Water from the West Spitsbergen Current (WSC), mixed with Arctic Water on the shelf. In addition to the variation in nutrient and salinity contributions from these overlapping water masses, the freshwater inputs from glaciers and rivers provide an important contribution to productivity rates here (Keck, 1999). As a result of these multivarious inputs, the fjord houses pelagic and benthic communities that comprise a mixture of boreal and Arctic flora and fauna (Hop et al., 2016).

Net primary production (NPP) in this region is controlled by a complex interplay of light and nutrients supplied by upwelling and lateral inflows from adjacent oceans and land which vary seasonally (Terhaar et al., 2021). Therefore, primary production is near-extant during polar night due to insufficient irradiance. As a result, one would expect very little abundance associated with light-harvesting phytoplankton at this time. However, field observations have demonstrated that virtually all taxonomic groups of Arctic microbes are present (Marquardt et al., 2016). This is demonstrated as winter protist communities characterized by extremely low abundance and biomass (primarily Bacillariophyceae, Ciliophora and Dinophyceae)(Kubiszyn et al., 2017). A shift in population dynamics/ species composition seems to be the common trend whereby heterotrophic eukaryotes (Syndiniales and Radiolarians) dominate in Winter, succeeded by a dominance of diatoms and the prymnesiophyte *Phaeocystis pouchetii* from Spring to post-bloom (Wietz et al., 2021).

3.0 Composition, utilization and biological effects of light

> *"Light is a general term which must be qualified in terms of "light climate" before it has meaning for biological systems."* (Berge et al., 2020)

Light climate and composition

Light climate can be defined as the *intensity*, *spectrum*, and *duration* of light for a given location (Berge et al., 2020). Each of these parameters can be measured/ defined in a variety of units, a

crucial aspect for their biological relevance as light-harvesting organisms are tuned to a specific light climate. When discussing it as a train of waves, it is described by its amplitude and wavelength (Brey, 2011). In the 'wave mode', different frequencies and wavelengths are associated with different colors. For this study, I will solely refer to the photosynthetically active part of the solar spectrum, comprising the area between ~400- ~700nm, often closely associated/ approximated to the visible light range ~380- ~740nm. Within this narrow range, we find wavelengths which are perfectly suited for absorption and utilization by light-harvesting organisms. PAR, discussed in terms of energetic yield, photosynthetic measurements, and how it is propagated through the ocean during polar night, is explored below in Box 1.

Box 1. Light Spectrum and PAR

Usage:
The various wavelengths within PAR are not utilized equally, this is due to their differential energetic yield. In the context of waves, different frequencies and wavelengths are associated with different colors. The lower the wavelength, the higher Photosynthetic photon flux density (PPFD). PAR is often likened to its associated quantum yield of CO_2 assimilation.

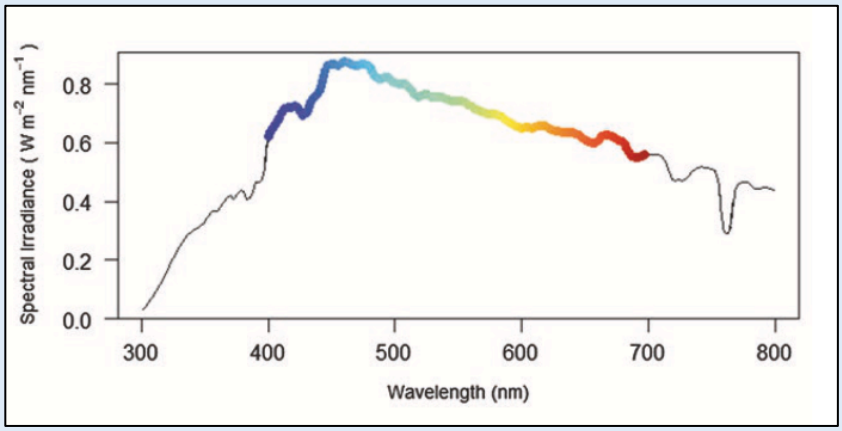


Figure 2. Spectral irradiance of Photosynthetically Active Radiation.

Energetic yield:

- Red and blue light are traditionally believed to have the highest quantum yield of CO_2 assimilation (*QY*, moles of CO_2 assimilated per mole of photons) and consequently are the most readily absorbed wavelengths by photosynthetic pigments (Liu & van Iersel, 2021). Despite blue-light ranging from ~380nm to ~500nm, it is one of the shortest, but highest-energy wavelengths.
- Light with a wavelength shorter than 400 nm or longer than 700 nm is considered irrelevant for photosynthesis due to its low quantum yield of CO_2 assimilation.
- However, the photosynthetic response of phototrophs to available PAR is not linear (Tripathy et al., 2014). Additional factors such as availability of well-suited machinery are equally as influential.

Photosynthetic measurements related to the photon scale:

- No matter the photosynthetic apparatus, the amount of energy captured from each photon will be the same, irrespective of their energy content.
- There will always be leftover energy, which will be dissipated as heat and Chlorophyll a (Chla) fluorescence (< 3%) (Sakshaug, E., Johnsen, G., and Kovacs, K., n.d.).

Underwater light in polar night:

is dependent upon:

- Atmospheric light sources (sun and moon).
- In-water light sources (bioluminescence).
- Optical properties of the water that influence these light sources.

Can be discussed in terms of "Inherent Optical Properties", the absorption and scattering properties of the water mass, influenced by particulate and/or dissolved constituents in the water. Data from Kongsfjorden suggest low IOP values that are homogenous with depth below the upper ~10m, reflecting a well-mixed water column during polar night (Berge et al. 2020).

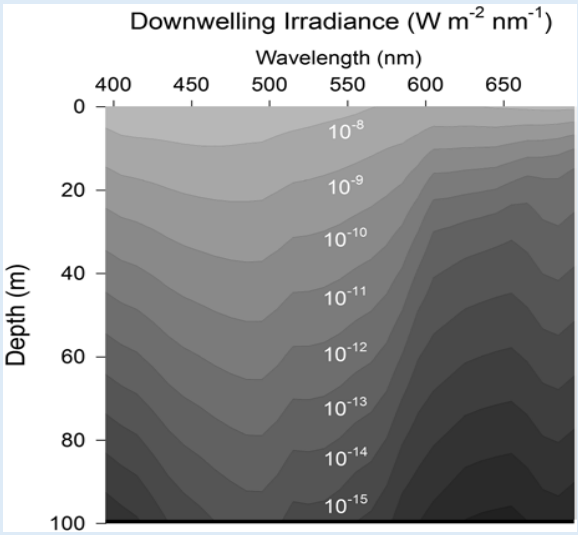


Figure 3. Modelled underwater spectral light in Kongsfjorden, Svalbard at midday under clear sky conditions, Berge et al. 2020.

Reference:
https://physics.stackexchange.com/questions/500659/how-to-shift-light-to-a-longer-wavelength-lower-frequency-energy-level-so-that
https://rwu.pressbooks.pub/webboceanography/chapter/6-5-light/

Light regime during polar night

The Arctic experiences extreme variations in both the magnitude and spectral composition of irradiance entering the ocean throughout the year (Connan-McGinty et al., 2022). During polar night, traditional definitions of day, night and seasonal photoperiod become irrelevant since there are only "twilight" periods defined by the sun's elevation below the horizon at midday (Cohen et al., 2021). However, solar elevation still significantly controls spectral irradiance during much of this period through atmospheric scattering of light from the sun.

Periods of clear sky during polar night show the blue part of the visible spectrum dominating, which is of high significance for blue to blue-green sensitive marine organisms who utilize this light for a range of biological processes (Båtnes et al., 2015; Valle et al., 2014). This causes us to review our recent view of polar night as being static and dark, to treating it as a dynamic photoperiod, characterized by different light regimes and sources, all of which are of biological relevance for photobiological studies.

The time period from polar night to spring can be associated with a in shift spectral irradiance (color and intensity), which is known to affect light-harvesting machinery in phytoplankton (Bercel & Kranz, 2022). The main difference to be accounted for between these two periods is the lack of sufficient PAR in the region spanning green-red light (~500-700nm) during polar night.

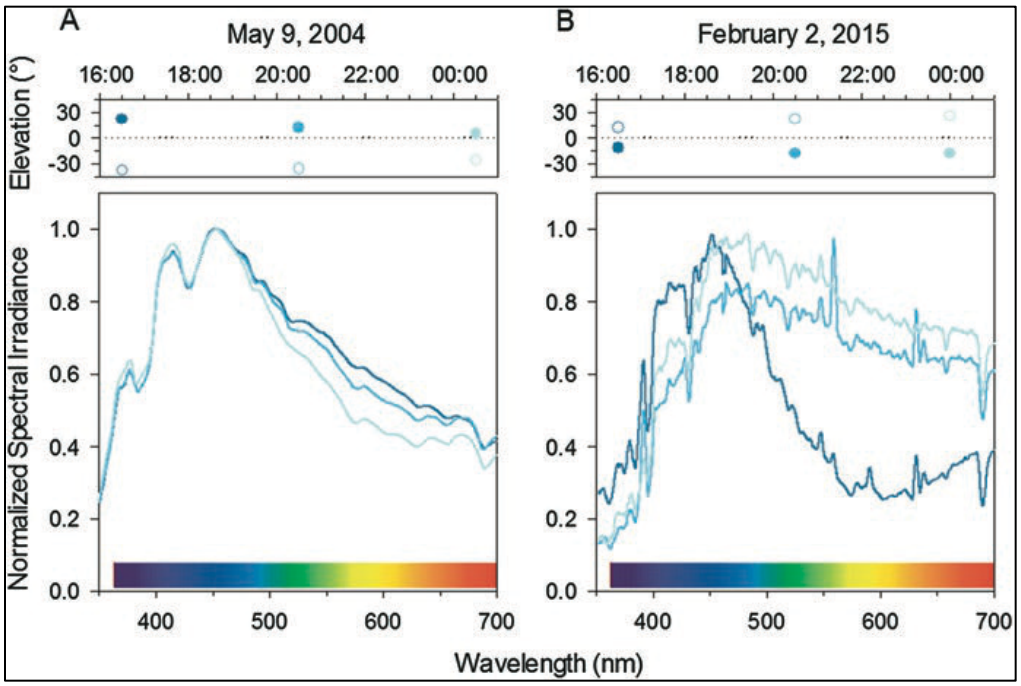


Figure 4. Changes in spectral irradiance when the sun or moon is near the horizon. Data are normalized irradiance at sea level in Ny-Ålesund and Figure from Berge et al 2020. A. Midnight Sun in May, when the sun (filled circles, upper panel) is always above the horizon and the moon (open circles) is always below the horizon. B. Polar night in February, when the sun is always above the horizon and the sun always below the horizon.

These changes can be understood by comparing irradiance when the sun or moon is near the horizon (Fig. 4) as it is critical for spectral composition and in turn biological utilization of light at any point during the 24 hour day (Berge et al., 2020; Palmer & Johnsen, 2015). At twilight with the sun below the horizon, a relative increase at blue and red wavelengths emerges as yellow wavelengths are selectively absorbed by atmospheric ozone (Cohen et al., 2021). Polar night and its associated moonlight can therefore be characterized by dim blue- and green-light whereas spring can be associated with white light (all wavelengths at equal intensity).

However, much remains unknown and/or not well understood with regards to the photobiological implications of the various sources of ambient light associated with polar night and how this may be manifested via light-harvesting gene expression. This study aims to contribute to the current knowledge gap by investigating a light regime in terms of spectral

intensity, composition and duration at an Arctic location, and its potential downstream ramifications on retinal-based photobiology.

Recent findings

Overwintering phytoplankton in the Arctic marine environment deploy various physiological adaptations and behaviours to survive the challenge of energy acquisition during polar night. These include utilization of stored energy products, formation of resting stages and reduction in metabolic rates (among others). However, we still lack comprehensive understanding of the status of light-harvesting machinery in-situ (with regards to prevalence, diversity, and transcriptional activity) in polar environments throughout polar night.

Many studies have investigated the effect of seasonality on gene expression in phytoplankton (Bowman et al., 2021; Diaz et al., 2023). An associated study at IsA demonstrated cell counts and biomass of microbial eukaryotes were lower during polar night compared to polar day (Kubiszyn et al., 2017). However, Operational Taxonomic Unit (OTU) diversity was inversely proportional to this trend, as has been previously reported in other Arctic and more-southern marine environments (Dasilva et al., 2014; Niemi et al., 2011). The study by Wutkowska et al., 2023 generated metatranscriptome-based knowledge regarding which marine microbial eukaryotes are present and active throughout the year in our study area (Svalbard archipelago). Their findings indicated that some key arctic microeukaryotic phototrophs were 'ribosomally active' during polar night (Wutkowska et al., 2023). Transcript abundance was higher during polar night than during polar day, by a factor of 2.7 (Wutkowska et al., 2023). The most abundant biological processes during polar day were related to respiratory electron transport chain or cytoplasmic translation. However, phototransduction mapping to blue- and green-light absorbing rhodopsin was also one of the most abundant Gene Ontology (GO) terms overrepresented in polar night. Phototransduction contained 208 transcripts mapping to green- and blue-light absorbing proteorhodopsins (Wutkowska et al., 2023). This unexpected result ignited the interest for this study and provided basal information for my search into rhodopsin abundance and expression in arctic phytoplankton.

4.0 Rhodopsins

Phylogenetic Origin

Rhodopsins are found in Archaea, Bacteria and Eukarya, and are diverse both genetically and functionally. Their initial discovery in the archaea *Halobacterium salinarium* was a major breakthrough in the understanding of energy acquisition in marine microbes (Béjà et al., 2000). The microbial rhodopsin family is comprised of more than 7000 photochemically reactive proteins in prokaryotes and lower eukaryotes found throughout the oceans from the tropics to the arctic (Govorunova et al., 2017). Based on their functions, they can be classified into light-driven ion pumps, light-activated signal transducers, and light-gated ion channels (Yoshizawa et al., 2022).

They comprise of seven transmembrane α-helices, with an all-*trans* retinal as the light-absorbing chromophore. Evidence indicates that their active site consists of at least three amino acid residues- Asp-85, Asp-212 and Arg-82- which are in close proximity to a protonated Schiff base linking the retinylidene chromophore to the transmembrane α-helices (Russell et al., 1997). The active site participates in proton transfers and regulates the visible absorption of bacteriorhodopsin and its photointermediates. As a result, each microbial rhodopsin exhibits a variety of specific visible absorption wavelengths of their retinal (Karasuyama et al., 2018), similarly to other accessory pigments (Table 1 and Fig. 5).

Table 1. Algal accessory pigments and their associated optimal absorption wavelengths.

Photochemical pigment	Absorption wavelength (nm)
Chlorophyll a	~372, ~642
Chlorophyll b	~392, ~626
Phycoerythrin	~495, ~545/566
Fucoxanthin	~510-525
Microbial rhodopsins	~483-489, ~515, ~532

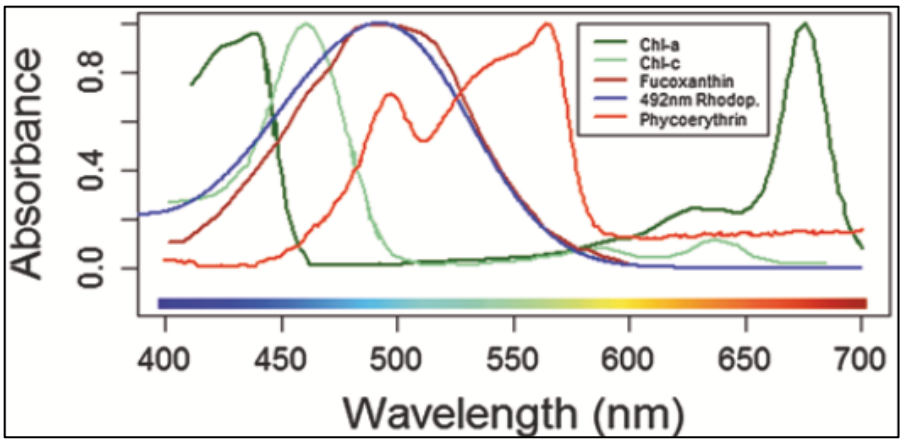


Figure 5. Absorbance of PAR by algal photosynthetic pigments (From Berge et al., 2020).

The proton-pump rhodopsin (PPR) absorbs light and drives protons across cell membranes, generating an outward proton gradient utilized for ATP production (Martinez et al., 2007). Thus, PPR enables cells to acquire energy from light, independently of plastid photosystems. These photoproteins seem to be spectrally tuned to absorb the wavelengths of light found in the surrounding environment: green light in surface waters and blue light at greater depths in the water columns (Shi et al., 2015). PPRs have been identified in some, but not all phytoplankton, with variability at the species level (Andrew et al., 2023). The presence of the conserved critical residues (retinal pocket, electron donor and acceptor) suggest that dinoflagellate rhodopsins likely arose through horizontal gene transfer due to their similarity to those found in bacteria (Fuhrman et al., 2008). However, a PPR gene is also found in several lineages of fungi and diatoms, where it is phylogenetically affiliated to the xanthorhodopsin clade, but seems to have a distinct evolutionary origin to that in dinoflagellates (Lin et al., 2010; Slamovits et al., 2011).

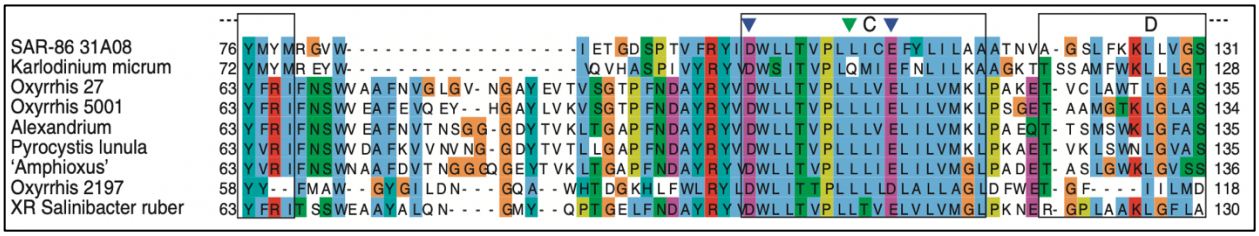


Figure 6. Amino-acid alignment of various rhodopsins from bacteria and dinoflagellates. Numbers on the right side indicate residue number. Black rectangles show predicted transmembrane segments. Functional sites (colored triangles): blue, proton acceptor, donor; green, spectral tuning (L for green and Q for blue) (Slamovits et al., 2011).

Figure 6 shows the sequence similarity between the functionally characterized proteorhodopsin SAR86-31A08 and various dinoflagellate species. Letters “D” and “E” (in purple, under blue triangles) and the proteins located in between them represent the active site. The single letter variation (marked by green arrow in Figure 6) depicted between SAR86-31A08 and *Karlodinium micrum* determines whether the rhodopsin will absorb in green-light (L for Lysine) or blue-light (Q for Glutamine). However, the functional extrapolation of bacterial PPR to dinoflagellate rhodopsin should be taken with caution due to the significant difference in cellular and molecular machinery between bacteria and eukaryotes (Shi et al., 2015).

Phytoplankton rhodopsin and relevance for Arctic strains

The current understanding of the prevalence and function of rhodopsins in marine phytoplankton remains scarce. Recent research based on transcriptomic analysis on lab cultures and natural assemblages shows that proton-pump rhodopsins are particularly widespread in different lineages of dinoflagellates. Examples of this are in *Oxyrrhis. marina*, *Prorocentrum* sp*., Polarella antarctica, Pyrocystis lunula, Alexandrium catenella* and *Karlodinium veneficum* (Guo et al., 2014; Lin et al., 2010; Meng et al., 2019; Ruiz-González & Marín, 2004; Shi et al., 2015; Stephens et al., 2020).

A study by Shi et al. 2015 investigating the regulation of rhodopsin gene expression in the Prorocentrales order saw that transcript abundance of the PPR gene exhibited a clear diel pattern with high abundance in the light period and low in the dark (Shi et al., 2015). Additionally, they observed that this rhythm was dampened when the cultures were shifted to continuous dark or light conditions, suggesting the gene is not under circadian clock control. Furthermore, expression levels demonstrated spectral tuning, with slightly higher transcript abundance under green as opposed to blue light. As these rhodopsin-promoting light conditions are similar to turbid marine habitats during a bloom formation, Shi et al. suggest that this gene may function to compensate for light-limited photosynthesis in dim environments. This concept is therefore highly applicable to the spectral conditions experienced by high Arctic phytoplankton during polar night.

Similarly, a study by Meng et al. 2019 investigated nonphotosynthetic photoenergy utilization by means of rhodopsins in two different strains of *Karlodinium veneficum*. Investigating such

mechanisms in strains isolated from coastal areas of the East China Sea, their findings suggest that both light intensity and light quality affected rhodopsin expression. The *K. veneficum* strains were found to respond differently to light conditions at photochemical and critical light-related genetic levels during both photosynthetic and nonphotosynthetic photoenergy utilization. Additionally, their observation of low-light adaptation in *K. veneficum* associated with rhodopsin expression is of great relevance when investigating this protein in high Arctic phytoplankton during polar night. The different spectral preferences exhibited by the two strains in the study demonstrates the challenge in applying similar methods to species from different geographical origin as genetic background and ambient nutrient conditions will affect expression patterns (Meng et al., 2019).

Very little is currently known on the abundance and expression of retinal-based photosystems in high Arctic phytoplankton, especially with regards to how this may be affected by light regime in a natural environment. A study published during the completion of this thesis investigated the widespread use of proton-pumping rhodopsin in Antarctic phytoplankton (Andrew et al., 2023). Their findings were that PPR is pervasive in Antarctic phytoplankton and localized to the vacuolar membrane in a model diatom. These unprecedented findings highlight the need for further research on rhodopsins in phytoplankton, specifically in polar habitats where the potential for their light-harvesting contributions are not yet recognized. As such, these papers were used for reference for methodology execution and later for data interpretation.

Research Goal

Experimental objective: Investigate the effects of seasonal changes in light regime on rhodopsin expression in high Arctic phytoplankton.

Research Aim 1: Describe the subsurface light climate in terms of intensity, spectrum and duration throughout the transition from the polar night to the spring equinox.

Research Aim 2: Design a qPCR assay for rhodopsin expression in high Arctic phytoplankton to determine rhodopsin abundance and expression during this time series

Research Aim 3: Explore environmental parameters as drivers of rhodopsin expression.

Materials and Methods

For this study, field campaigns were set up to form a time series spanning the transitory period from polar night to the spring equinox. During field campaigns, environmental data was collected (CTD, nutrients, spectrofluorometer) as well as water samples for DNA and RNA extracts for use in targeting natural assemblages of dinoflagellates and diatoms for rhodopsin gene abundance/ expression. In order to investigate the time series environmental samples, a qPCR assay for rhodopsin expression in high Arctic phytoplankton was designed. This was done via the use of primers targeting various light-harvesting genes (blue- and green-light absorbing rhodopsins, and photosynthetic genes) either designed or copied from other papers (Levialdi Ghiron et al., 2008; Meng et al., 2019; Morgan-Kiss et al., 2016). These underwent extensive testing via PCR in the hopes of isolating the rhodopsin-encoding region. Following the optimization of primer design (bacterial cloning and Sanger sequencing) and reaction parameters (temperature trials, nested PCR, qPCR) the qPCR assay was used on the time series samples to investigate the transcriptional activity of rhodopsin in two high Artic strains of phytoplankton and how this may be a function of ambient spectral and environmental drivers.

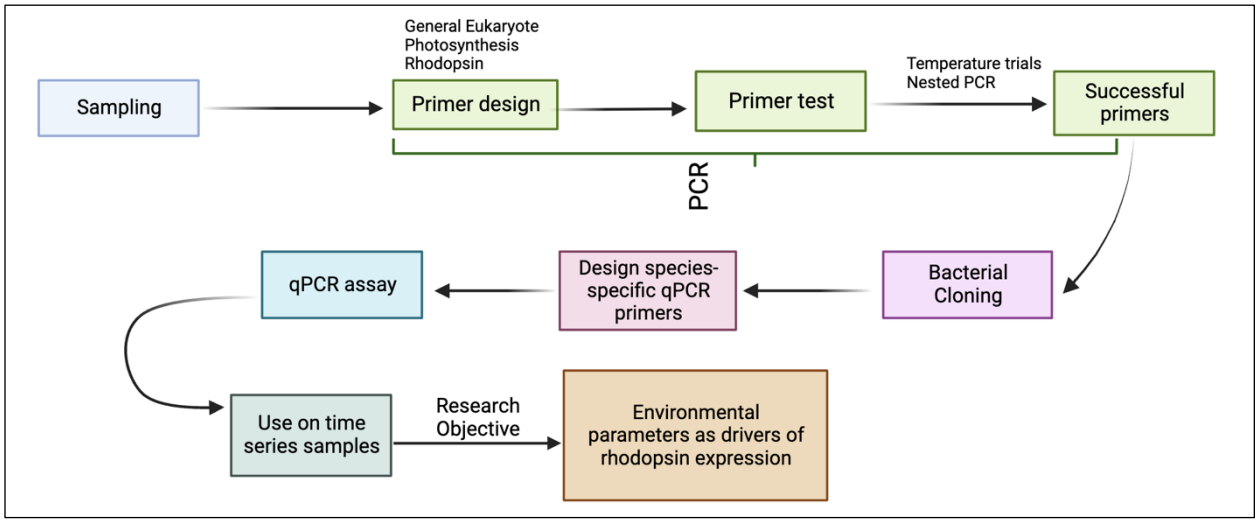


Figure 7. Execution of methodology.

1.0 Sampling

Three sampling periods occurred for this study, each taking place within a separate light climate.

The first took place during a polar night cruise in January 2021, and the second in September 2022, when daylength exceeded 14.5 hours. The sample from the latter was collected using a 12L bucket lowered over the side of a polarcirkel boat at the IsA time series station (78°15'.669''N, 15°32'.023''E) to 1m below sea level. The DNA and RNA from these field campaigns were to be used as test material for the first set of primers.
The third sampling period from January 2023 to April 2023 compiled a time series spanning the transition phase from polar night to spring equinox- when the light cycle is 12 hours light, and 12 hours darkness. All samples were collected as close to local noon as possible (10:30-11:30a.m CET). An SD208 CTD probe from SAIVA/S (Bergen, Norway) was deployed at each field campaign in the time series. Despite this study solely investigating light-harvesting activity in phytoplankton at the sea surface, the instruments were deployed to a maximum depth of 85-90m to gain a vertical profile of ambient conditions in the water column throughout the time series. Save for the first field campaign, a spectrofluorometer (Photosynthetically Active Radiation sensor by In-situ Marine Optics) containing a Si photodiode sensor was also deployed to study spectral irradiance in the water column. Only data from 1m below the sea surface will be shown from this instrument.

Time series sampling procedure

Samples were collected twice a month (see Table 2 for information) and always in the same way: a 10 L Niskin bottle was lowered over the side of the vessel Polarsyssel, to a depth of 1m below the surface. Of this, two clean 50mL bottles were taken as subsamples for nutrient analysis at a later date. The remainder of the sample was divided into two separate 4L Nalgene containers (one for DNA and one for RNA) and kept in a cool box during transport. Once at UNIS, the falcon tubes destined for nutrient analysis were kept at -20°C until further use.

Table 2. Field campaigns throughout time series.

Sampling date	Sample purpose
15/01/2021	Test material
14/09/2022	Test material
17/01/2023	Time series
01/02/2023	Time series
15/02/2023	Time series
01/03/2023	Time series
15/03/2023	Time series
31/03/2023	Time series

2.0 Sample processing

All protocols mentioned below can be found in full in Supplementary materials.
Keeping DNA and RNA samples separate, samples were poured through a 65um mesh (KC Denmark) to remove larger zooplankton, collecting target organisms on a Durapore® 0.45µm PVDF filter (Millipore) using a vacuum filtration pump supporting three 250mL funnels. This resulted in 1x DNA filter and 1x RNA filter per field campaign. DNA filters were cut in half on a sterile petri dish, placed in two 2 mL cryo tubes and stored at -80°C until extraction. RNA filters were placed in 2 mL cryo tubes with 600μl 6 Lysis/Binding Solution (RNAqueous kit, Thermo Scientific™) and stored at -80°C until extraction.

2.1 DNA Extractions

Total DNA was extracted from the filter halves using the DNeasy® Plant Kit by QIAGEN, according to the manufacturer's recommendation with an extra bead beating step (protocol in supplementary). An extraction blank was included to ensure that the extraction had worked. DNA was eluted with 150μl elution buffer per filter half, resulting in 300μl of elute with DNA per sampling date. After extraction evaluation, the replicates of each field campaign (filter halves) were added together.

The concentration was evaluated using an Invitrogen™ Qubit™ 4 Fluorometer (Qubit 1X dsDNA HS Assay Kit according to manufacturer's recommendations), of which the protocol can be found in supplementary. The quality of the extraction was evaluated by performing PCR using general eukaryotic primers (Piredda et al., 2017) to test for contamination. Briefly, 1μl of DNA extract was combined in a PCR strip with a 24μl mastermix containing 2.5μl DreamTaq Buffer, 0.25μl DreamTaq DNA Polymerase (both from ThermoScientific™), 2μl dNTP mix (2.5 mM), 0.5μl each of forward and reverse primers (10 µM) from Piredda et al., 2017 and 18.5μl Milli-Q H_2O. The PCR program can be found in supplementary. 5μl of post-PCR DNA was run on a 1.0% agarose gel in 1x TAE buffer, with 3 μl low range ladder (Thermo Scientific ™) was used for size indication. The extraction blanks did not produce any product, indicating a successful extraction without signs of contamination. Working aliquots of 50μl were made from these extracts and stored at -20°C to prevent continuous freezing and thawing. The remainder of the extracts were kept at 80°C until further use.

2.2 RNA Extractions

Total RNA was extracted with the RNAqueous Total RNA Isolation Kit (Invitrogen, Thermo Scientific™), according to the manufacturer's recommendation, with an extra bead beating step (protocol in supplementary). A DNAse treatment step was performed on all extracted RNA, with the inclusion of an extraction blank to ensure successful extraction, using the TURBO DNA-*free*™ Kit by Thermo Scientific™, with the protocol provided in the kit. The DNAse treatment removes trace amounts of contaminating DNA by enzymatic nonspecific cleavage of DNA. Briefly, 1 μl of RNA was combined either with 1 μl random hexamer primer (100 µM, Thermo Scientific™), which is unspecific and will bind anywhere, or with 1 μl a poly-A tail-specific primer ('OP41',10 µM) to see if exclusively eukaryotic transcripts could be targeted (Vader et al., 1999). These will hereinafter be referred to as *RH* cDNA and poly-A cDNA, respectively. Both were combined with 1 μl dNTP mix (2.5 µM) and 10 μl nuclease-free water (Ambion) and denatured by incubation in a PCR machine (Eppendorf Mastercycler® x50) for 5 minutes and immediately placed on ice. Subsequently, cDNA synthesis was performed using the Superscript ™ protocol by Thermo Scientific™ on PCR strips. 13 μl of denatured DNAse treated RNA was combined with 4 μl 5x Superscript IV buffer, 1 μl Superscript IV RT (200U/µl), 1 μl 0.1M DTT

and 1 μl RNAse inhibitor (40U/µl, all of which Thermo Scientific™), and denatured by incubation in PCR machine (refer to supplementary for program). All cDNA synthesis (polar night, September sampling and time series RNA) was tested by via PCR, using general eukaryotic primers from Piredda et al., 2017 (same materials, quantities and program as those mentioned above for evaluation of DNA extraction). However, the use of poly-A cDNA would later be limited to primers targeting nucleolar transcripts as the poly-A tail is a characteristic trait of messenger RNA of nucleolar origin.

Quality of the RNA extraction after DNAse treatment and cDNA synthesis was done using agarose gel electrophoresis, whereas quantity was determined using a Qubit™ (Qubit RNA HS Assay Kit according to manufacturer's recommendations). The instrument displayed "Out of Range" for one sample (RNA on 01/03/2023).

Table 3. Extraction concentrations obtained using Qubit™ Fluorometer. Replicates of 2 for DNA is due to filters being cut in half and extracted/eluted separately.

Sample ID	Filter number	DNA ng/ul	RNA ng/ul
17/01/2023	1	7.36	4.00
	2	6.78	/
01/02/2023	1	3.82	3.05
	2	2.04	/
15/02/2023	1	2.76	3.05
	2	2.34	/
01/03/2023	1	2.08	NA
	2	2.64	/
15/03/2023	1	1.42	2.40
	2	1.73	/
31/03/2023	1	2.18	2.62
	2	3.38	/

The evaluation using gel electrophoresis was done on a 0.7% agarose gel in 1x TAE buffer on a unit exclusively reserved for RNA samples. Samples were loaded into the gel such that for each field campaign, one well was loaded with untreated RNA (pre-DNAse), the next with DNAse treated RNA, and the last with the sample from cDNA synthesis. 3 µl low range ladder (ThermoScientific ™) was used for size indication. Both untreated and treated RNA failed to

produce any amplification products, showing that DNAse treatment had been efficient and that the amplification product in the cDNA reactions was due to reverse transcribed RNA. Working aliquots of 50μl were made from these extracts and stored at -20°C to prevent continuous freezing and thawing. The remainder of the extracts were kept at 80°C until further use.

2.3 Nutrient Analysis

Nutrient analysis was done using a QuAAtro 39 Nutrient Analyzer from SEAL Analytics™, according to the manufacturer's recommendations. The samples (2x 150mL falcon tubes per field campaign) and analyzed for Silicate (SiO_4), Nitrate (NO_3) and Nitrite (NO_2), Phosphate (PO_4). Nitrate and Nitrite values were combined to display recordings as "NO2+3" or NOx. The limits for detection and quantification can be seen in Table 4. Negative values should be considered as zero as long as they are -1.5 - 0.0.

Table 4. Limits of detection and quantification for Nutrient Analyzer.

	SiO_2 µmol/L	NO_3^- µmol/L	PO_4 µmol/L	NO_2 µmol/L
Limit of detection (LOD)	0,02	0,05	0,02	0,01
Limit of quantification (LOQ)	0,04	0,07	0,08	0,04

3.0 Primer design

In order to explain the methodological route taken for primer design, it is important to first outline the current state of understanding and sequence availability for rhodopsin in phytoplankton and how these can be applied to high Arctic strains.

So far no such attempt has been made to investigate rhodopsin in high Arctic phytoplankton. As such, sequences from more temperate latitudes -East China Sea (Ma et al., 2023; Shi et al., 2015) and Long Island Sound (H. Zhang et al., 2022)- as well from an associated study (Wutkowska et al., 2023) were used as references for blue- and green-light absorbing rhodopsin. Primer design began on a broad spectrum (Fig. 8) as I was unsure of the conserved status of these sequences

across phyla and geographic locations. Later, optimized primer design would enable me to streamline my search, resulting in primers designed for rhodopsin in specific strains of high Arctic phytoplankton for use in qPCR.

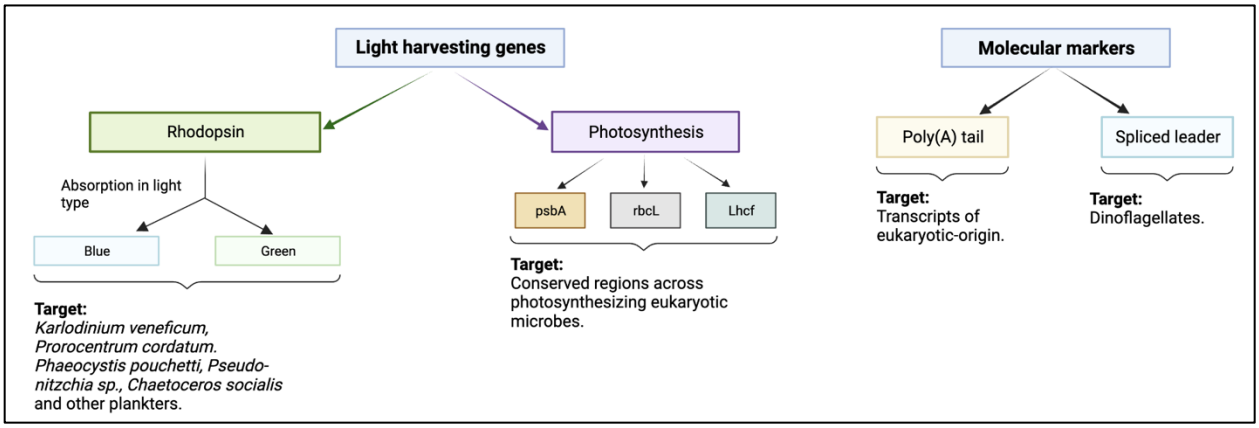


Figure 8. Flowchart of target regions for primer design.

Therefore, I originally designed many primers for blue- and green-light absorbing rhodopsins in phytoplankton species which are usually present in the Arctic, and whose rhodopsin sequences are known (*Karlodinium veneficum, Prorocentrum cordatum. Phaeocystis pouchetti, Pseudo-nitzschia* sp*., Chaetoceros socialis*). I also chose to integrate the study of photosynthesis-related genes in general marine eukaryotes to see if they could be used as a control/ proxy for light-harvesting transcriptional activity. I chose to utilize certain eukaryotic molecular features to ensure that the targeted genes were of eukaryotic origin and not from their bacterial counterparts. Those being the poly-A tail in eukaryotic transcripts, and the spliced-leader in dinoflagellates.

3.1 Designing primers from sequence alignments

SeaView,Version 5.0.4, (Gouy et al., 2010) was used to align and visualize the reference rhodopsin sequences from literature, the Wutkowska et al., 2023 study (Fig. 9), and National Centre for Biotechnology Information (NCBI). Once aligned, areas of similarity amongst the sequences were utilized for primer design. Viewing them in this manner also enabled me to use the Basic Local Alignment Search Tool (BLAST) from NCBI to find regions of local similarity

between nucleotide or protein sequences, and later for primer BLASTs to verify their similarity to other sequences and likelihood to form primer dimers.

Box 2. Primer Design

Primer properties:

- Specificity depends on primer length and annealing temperature.
- T_m: Melting Temperature, where one half of the DNA duplex will dissociate to become single stranded. The optimal melting temperature for maintenance of primer specificity is generally in the 52-58°C range.
- Primer pairs should not have T_ms with a temperature difference >5°C.
- During a PCR, T_m describes the temperature at which primers have annealed to 50% of the target sequences and the other 50% of target sequences are free, hence, equilibrium has been reached.

Other factors affecting reaction success:

- Primer concentration
- Buffer composition
- Metal ion concentration,
- pH and DMSO within the PCR mastermix[1].

Degenerate primers:

- Designing primers in a 'degenerate' manner, meaning a mixture of similar primers with different bases at the variable positions, is a solution if the exact nucleotide sequence of the target region is unknown. They are designed by aligning homologous genes or amino acid sequences from multiple species, making it possible to visualize conserved or non-conserved regions between species.
- The **benefits** of designing primers in this way is the increased chances of compatibility with transcripts by designing single nucleotide differences, giving a population of primers with similar sequences that cover all possible nucleotide combinations.
- The **downside** of using degenerate primers is that it can lead to difficulties in optimizing PCR assays. Only a limited number of primer molecules are complementary to the template, and the melting temperature (*T*m) of primer sequences vary significantly depending on the nucleotide type content (e.g high GC content raises the *T*m) therefore making it difficult to not melt the primers during amplification.

Optimized primer design:

- Target area should be approximately 75-300 base pairs in length for optimal PCR amplification.
- The less degenerate the primers are, the further apart they can be.
- GC content between 40-60%.
- Include between 6-7 amino acids in the primers (equal to 15-20 nucleotide base pairs).
- Avoid degeneracy in the 3 nucleotides at 3' end.
- 3' end should terminate in G or C to promote binding.
    - G and C bases have stronger hydrogen bonding and help promote stability.
- Refer to *T*m Calculator by *New England BioLabs ®Inc* to know whether the primers are outside of the optimal temperature difference range + recommended annealing temperature.

Table 5. IUPAC degenerate primer nucleotide

IUPAC nucleotide code	Base
A	Adenine
C	Cytosine
G	Guanine
T (or U)	Thymine (or Uracil)
R	A or G
Y	C or T
S	G or C
W	A or T
K	G or T
M	A or C
B	C or G or T
D	A or G or T
H	A or C or T
V	A or C or G
N	Any base

References:

1. https://www.qiagen.com/us/knowledge-and-support/knowledge-hub/bench-guide/pcr/introduction/guidelines-for-degenerate-primer-design-and-use
2. https://bitesizebio.com/18992/a-primer-for-designing-degenerate-primers/#:~:text=1)%20Align%20multiple%20amino%20acid,further%20apart%20these%20can%20be.

As such many constraining conditions had to be considered simultaneously during primer design (Box 2). I aimed for lengths of 18-30 nucleic bases, a target area of 75-200bp, minimal difference between their melting temperatures (when feasible), and a GC content of 50%. As the conserved status of these genes across phytoplankton genera is still unknown, many degenerate primers were designed for several regions in the hopes of having hits on similar sequences in more than one species. Using degenerate primers in the early stages enabled me to target gene sequences whilst only possessing the associated sequence in related/ similar organisms. Through a process of elimination, this methodology would enable the search for rhodopsin-encoding regions to be streamlined for specific high Arctic phytoplankton strains. All primers which were designed can be found in tables S2-S5 in supplementary material.

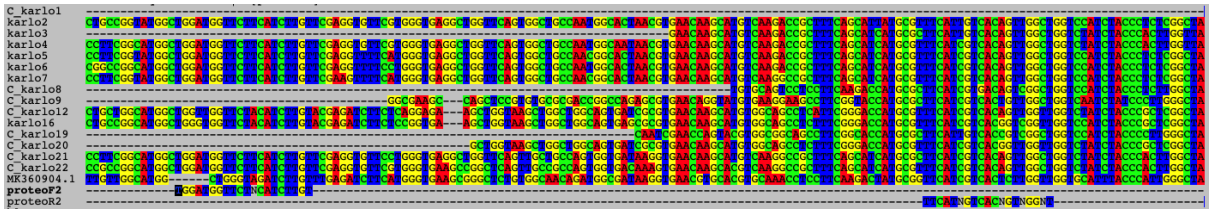


Figure 9.A. Karlodinium rhodopsin nucleic acid sequences with associated primers (proteoF2, proteoR2).

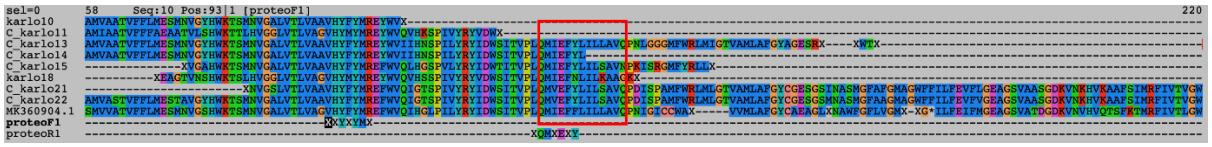


Figure 9.B. Karlodinium rhodopsin protein sequences with associated primers (proteoF1, proteoR1) and active site region from letters "D" to "E" (red box).

Figure 9.A can be referred to for demonstration of primer pair *proteoF2*/*proteoR2* in relation to their target nucleotide region within *K. veneficum* rhodopsin sequences from the Wutkowska et al. study. Sequences were also aligned as proteins (Fig. 9.B) to view the Q and S active site of proteorhodopsin and ensure that some primers were upstream of this region, with the hopes that amplicons would contain this site.

3.2 General eukaryotic rhodopsin primers

Rhodopsin-targeting primers were designed continuously throughout this study, each time optimizing the design process and narrowing down the target area.
Blue-light absorbing rhodopsins (proteorhodopsins) were targeted using sequences from the Wutkowska et al. study which had: a) a hit to *K. veneficum,* or b) a hit to uncultured eukaryotes which are associated with the Q and S active sites proteorhodopsins. The target regions for the degenerate primer pairs ranged from ~70-100bp whereby many forward and reverse primer combinations were possible. Green-light absorbing rhodopsins (xanthorhodopsins) were targeted using sequences from the Wutkowska et al. study which had: a) hits to *Prorocentrum* sp., or b) hits to *Phaeocystis pouchetii*.

3.3 General eukaryotic photosynthesis primers

Various photosynthesis-related primers were used with the hopes of investigating light-harvesting activity under changing spectral conditions, specifically in comparison to rhodopsin-encoding regions. Studies investigating light-harvesting activity in phytoplankton (via photosynthetic and retinal-based photosystems) were referred to for sequences (Levialdi Ghiron et al., 2008; Meng et al., 2019; Morgan-Kiss et al., 2016; Y. Zhang et al., 2019). These sequences were chosen as they are known to be conserved across various phyla.
As such, I hoped that their use could provide a proxy for expression in light-harvesting machinery in eukaryotic phytoplankton. The photosynthetic genes were the following: 1. RbcL, a chloroplast gene encoding the large subunit of ribulose-bisphosphate carboxylase, therefore only to be tested with random-hexamer-cDNA, 2. Lhcf: Light harvesting gene family and 3. psbA gene family encoding the D1 protein of photosystem II, both of which were tested with *RH*-cDNA and poly-A cDNA.

3.4 Primers targeting the poly-A tail and dinoflagellate spliced leader

Poly(A) tails are a eukaryotic transcriptional trait which are added to the majority of mRNAs during the 3’ end processing stages in a co-transcriptional fashion. Despite their simple sequence composition, 3’ terminal polyadenosine tracts play critical roles in multiple aspects of a transcript’s life cycle (Jalkanen et al., 2014). Their roles span facilitating exonuclear

translocation, to key regulatory roles in enhancing translation efficiency. The length of the poly-A tail varies greatly, however it is believed that most are 150-250nt (Jalkanen et al., 2014). It was therefore decided to use a poly-A tail-targeting primer from a previous study to isolate eukaryotic transcripts from the samples. Primer OP41, ensures the capturing of the poly-A tail due to the complementary Thymine bases (Vader et al., 1999). When used in combination with complementary primer OP283 (from same study), eukaryotic transcripts with poly-A tails can successfully be amplified.

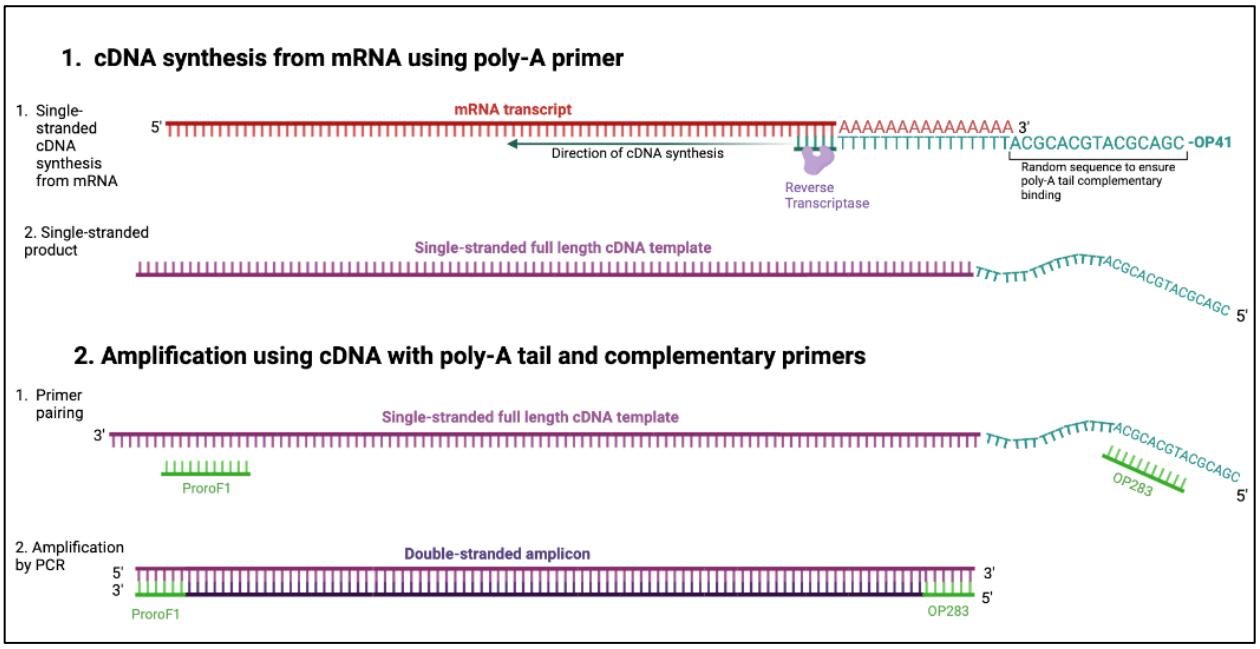


Figure 10. cDNA synthesis from mRNA using poly-A primer OP41 and further amplification.

Composed of a short RNA fragment (~15-50nt), the spliced-leader can be found at the 5'-end of nucleus-encoded mRNAs in dinoflagellates whereby it is transferred from the 5'-end of a small non-coding RNA (SL-donor transcripts). The conserved sequence is: "5'-DCCGUAGCCAUUUUGGCUCAAG-3'" (Zhang et al., 2009). A spliced-leader targeting primer was designed (*dinoSL*) to be used in combination with various reverse blue- and green-light absorbing rhodopsin reverse primers (Rps). This was done to target gene expression for a natural dinoflagellate assemblage amidst other co-existing organisms.

3.5 Species-specific primers

It was then decided to narrow my area of research by choosing 5 phytoplankton species to investigate rhodopsin expression in (*Karlodinium venificum, Phaeocystis pouchetii, Chaetoceros socialis, Prorocentrum* sp.*,* and *Pseudo-nitzschia*). Species-specific primers targeting the following genes were designed: Rhodopsins, rbcL, lhcf and psbA protein. This was done using available sequences from NCBI. Additionally, primer sequences from Zhang et al., 2019 study which, upon metatranscriptome analysis of a *Prorocentrum donghaiense* bloom showed evidence of as many as 375 active rhodopsin genes (Y. Zhang et al., 2019), were used to see if rhodopsin sequences were conserved in our high arctic strain of *Prorocentrum* sp.. Their target regions varied in size, with potential amplicons being ~78, ~189 and ~226bp.

Eventually, a final set of 26 primers were designed solely for the qPCR assay, specifically targeting rhodopsin in *Karlodinium veneficum* and *Prorocentrum* sp. (Table 6).
Several primer pairs were designed per species, both degenerate and not, and including the active site for rhodopsin and not. This was done to be able to experiment with Fp/Rp combinations to optimize reaction parameters for the assay.

4.0 Primer testing

All primers designed for this study underwent thorough testing via PCR before use in the qPCR assay. Unless specified, programs and tests can be found in supplementary material. All PCR mastermixes were set up in a fume hood, and PCR strips were set up on a sterilized bench in a sterile lab.

4.1 Reaction parameters

The mastermix (M.m) comprised of DreamTaq DNA Polymerase (5 U/ µl), 10X DreamTaq Buffer and Nuclease-free water from Thermo Scientific™, dNTP mix (10µM) and forward and reverse primers diluted to 10 µM. Once prepared in 1.5 mL Eppendorf™ tubes, the M.m was spun down in a microcentrifuge (VWR® MiniStar silverline microcentrifuge) and kept in a Mini-cooler rack until use. For most primer pairs, the recommended Master mix (M.m) to

cDNA/DNA ratio (24µl M.m : 1 µl cDNA/DNA/ Milli-Q H_2O for negatives) was used. When amplifying smaller target regions (i.e 72 bp), was often augmented to 2 µl cDNA/DNA. PCRs were set up in MicroAmp™ strips to which M.m was added to the genetic material/Milli-Q H_2O and mixed by pipetting. PCRs were performed on an Eppendorf Mastercycler® x50 with the following program: 94°C for 2', 30 cycles at 94°C for 30", 45°C for 30", 72°C for 1', incubation at 72°C for 7', holding phase at 10°C. The cycle number was generally kept at 30x as increasing this to 35x could lead to an increase in unspecific products.

4.2 PCR trials.

The total of 72 primer pairs (117 individual primers with various combination possibilities) both designed for this study as well as those from literature were tested at different time periods depending on the target gene, species, and taxonomic grouping (see supplementary Tables S7-10 for primer combinations and temperatures tested). Throughout the +65 PCRs performed towards the development of the qPCR assay, the following criteria were used to determine if a primer pair had amplified successfully:

1. No or little primer dimer formation.
2. No amplification in the negative.
3. Product of expected size when verified on agarose gel.
4. Amplification with different genetic material.

The first stage of testing involved 32 primer pairs targeting: 1. Blue- and green- absorbing rhodopsins and 2. Photosynthesis genes, 3. General eukaryotes (Godhe et al., 2008; Huang et al., 2019; Piredda et al., 2017), using *RH* cDNA from both the polar night and September field campaigns. All primer pairs were originally tested at 50°C, with 1 negative per M.m/ template replicate. For primer pairs with promising results observed on agarose gel, a temperature trial was performed to find optimal reaction temperatures and parameters. If results indicated signs of contamination, new primer aliquots were made (10 pmol/ µl), and dNTP mix, DreamTaq, and/or 10x DreamTaq Buffer were replaced.

Eventually, time series DNA and RNA (*RH* cDNA and poly-A cDNA) was tested with successful primers from previous testing stage and the 20 newly designed species-specific (*Karlodinium. veneficum, Prorocentrum* sp*., Chaetoceros socialis, Phaeocystis pouchetii* and *Pseudo Nitzschia* sp.) rhodopsin-targeting primers. Once optimal results were obtained, chosen primer pairs were used for nested PCR, bacterial cloning and Sanger sequencing.

4.3 Temperature trials.

The use of degenerate primers makes it difficult to determine the annealing temperature for PCR. Therefore, for promising primer pairs and those who had a T_m with a difference of 5°C or more, temperature trials were performed. Temperature trials were performed spanning several temperature ranges, i.e: 48-52°C, 50-55°C, 53-58°C. These were later verified on agarose gels to test for primer dimers and stronger bands at certain temperatures. On several occasions the difference in T_m between two primers was well above the recommended 5°C range, meaning that a temperature trial covering a range of 5-10°C difference was necessary to conduct. This makes the pair very prone to denaturation once the reaction temperature gets too high. Primer pairs were subjected to temperature trials on various versions of cDNA in attempt to find optimal reaction parameters.

4.4 Nested Polymerase Chain Reaction.

To increase the specificity of amplicons for use in bacterial cloning, I performed nested Polymerase Chain Reactions. In this, the product of the first amplification reaction is used as the template for the second PCR, and both sets of primers have different and unique properties. As my rhodopsin-targeting primers are based on the notion that these sequences may be conserved across phytoplankton genera, using nested PCR enabled me to minimize areas of uncertainty within the transcripts. The ‘outer primers’ are those used for the first PCR and the ‘inner primers’ are those for the second. The outer primers are upstream of the inner primers and are bound to the outside region of the target DNA and amplify larger fragments (poly-A tail included). If non-specific DNA sequences are amplified in the first round of PCR, they will not be amplified further in the second PCR.

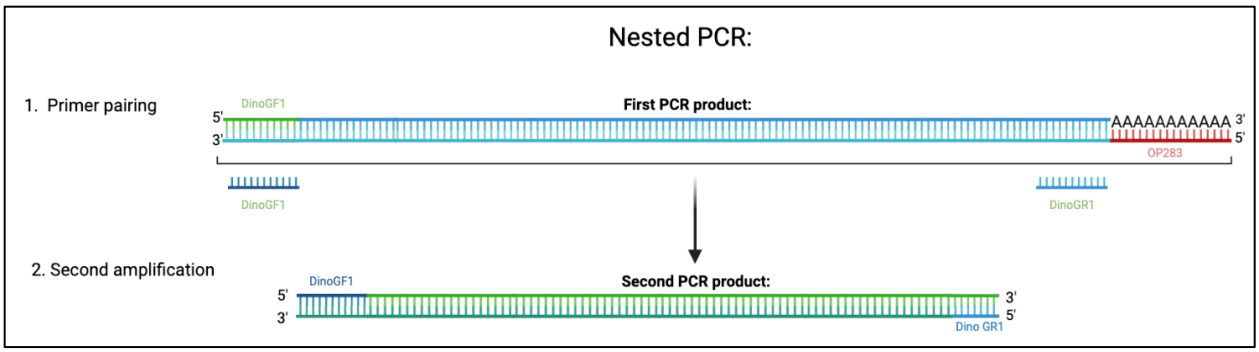


Figure 11. Nested Polymerase Chain Reaction with Dinoflagellate and poly-A tail primers.

For the first PCR, I chose a forward primer which had demonstrated a high success rate and amplified it paired with a poly-A reverse primer (Rp)- OP283- resulting in unspecific products, usually larger than the target amplicon size. This reaction was done at 50°C using poly-A cDNA from polar night. Doing so enabled me to refine my target region and remove unwanted amplified regions upstream of the forward primer, whilst ensuring amplicons of eukaryotic origin.

For the second PCR, the Fps' associated Rps bound specifically at the target site and amplified the desired sequence. To increase specificity, the temperature was raised to 52°C from the normal program, however the cycle number was kept at 30 as a first attempt at nested PCR with 35 cycles led to multiple different sized products.

4.5 Bacterial Cloning

The nested PCR products used for cloning were the following primer pair combinations: *BlueF/ProteoR2, DinoGF1/DinoGR1, PhaeoF1/PhaeoR1* and *RhbF2/RhbR2* and were chosen based on results from Sanger sequencing which revealed hits to *Karlodinium venificum* and *Prorocentrum* sp. All samples were cleaned using SPRI beads to remove leftover primers from the products. The PCRs were then repeated to have freshly made products for cloning.
An overnight culture of *Escherichia coli* was prepared, and plates made using Luria-Bertani (LB) liquid medium. Ampicillin was later added to reach a final concentration of 50µg/ml.
The ligation of the PCR products was performed using 5ng of product as the lengths of the amplicons were ~300, ~120, and ~79bp, respectively.

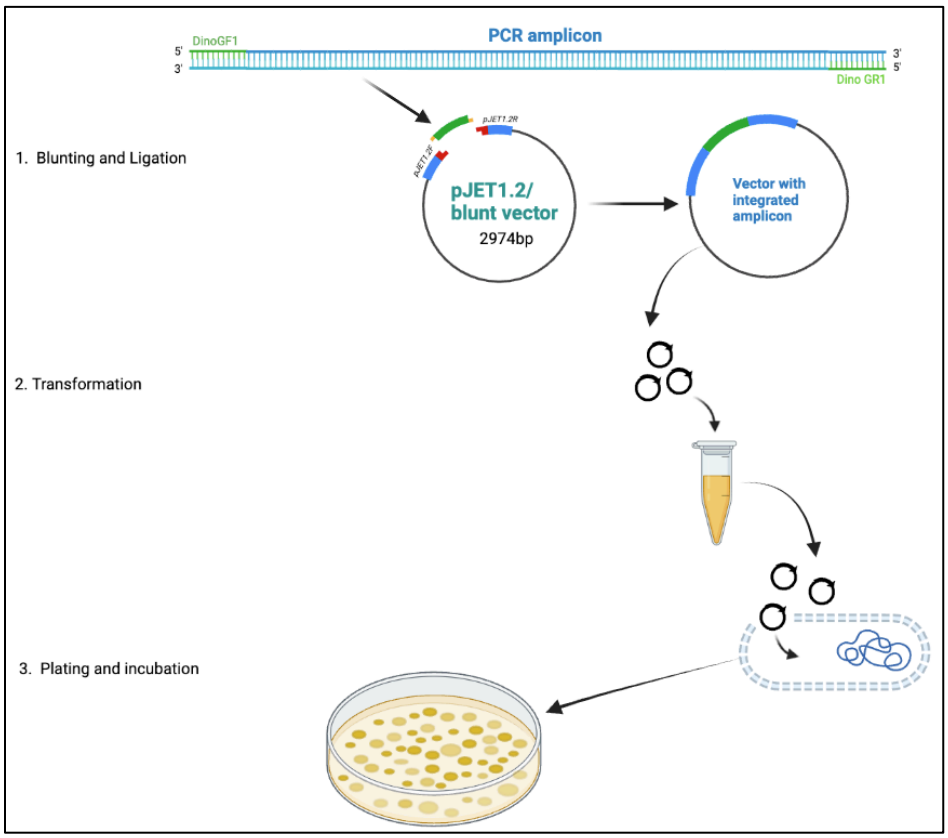


Figure 12. Bacterial cloning of PCR amplicons.

Ligation was performed using the CloneJet PCR Cloning Kit and the TransformAid Bacterial Transformation Kit, both from Thermo Scientific™.

Once ligation and transformations had been performed, the products were transferred to individual LB plates. Separate plates were prepared for both negative (no genetic material) and positive (pUC19 vector from Thermo Scientific™, 2 µL of pUC19 10 pg/µL) controls. The positive control using the pUC19 cloning vector provides a measure of transformation efficiency, whilst the negative control provides indication of contamination or human error during LB medium preparation, ligation or transformation. The plates were incubated at 37°C overnight. To test whether the colonies have inserts of the expected size, single colonies from each plate were then transferred to a grid colony plate using a clean pipette tip. This was once again incubated overnight at 37°C. Colony PCR was performed using the forward and reverse primers provided in the kit (pJET1.2F & pJET1.2R) on a total of 26 bacterial colonies to ensure that the target region is within the vector where the amplicons were inserted. The PCR products were then

checked on a 1.0% agarose gel (1x TAE) revealing that the product sizes fit the expected length, and they were sent for Sanger sequencing. A second repetition of cloning was performed in the hopes of obtaining more colonies. This deemed more successful therefore the colony PCR products were sent for Sanger sequencing. Bacterial cloning was later repeated for preparation of the standard curve for qPCR.

4.6 Final stage

The final stage involved testing the primer pairs which were designed from Sanger sequencing cloning results indicating a hit (NCBI genebank) to *Karlodinium veneficum* and *Prorocentrum donghaiense* rhodopsin sequences (Table 9), using time series poly-A and *Random Hexamer* cDNA, as well as time series DNA. The 3 most promising primer pairs were selected for the qPCR assay (Table 6) as they desmonstrated high success rates with environmental samples (verified via agarose gel electrophoresis indicating crisp bands and lack of primer dimers) and Sanger Sequencing results.

Table 6. Primers pairs chosen for qPCR assay.

Primer pair	Sequence (5'-3')	Expected product size (bp)
KarloF4/R4	TGCAGATGATCGAGYTCTA CACCCGCGAARATCTCGA	~210
KarlorhodF2/R4	TCCTCTCTGCAGTGCAGC CCTTCRCCCGCGAARATCTC	~180
ProroF1/R1	AGATAACGACGCATACCG GTAGCCAGAGATGATCAT	~160

Two primer pairs targeting *K. veneficum* and one targeting *Prorocentrum* sp. were chosen. The Proro F1/R1 primer pair targets a region which includes the active site for *Prorocentrum donghaiense* rhodopsin. However, designing similar primers for *K. veneficum* proved difficult due to a lack of sequence homology, hence the inclusion of two primer pairs to expand my scope.

5.0 qPCR

5.1 qPCR preliminary testing

qPCR trials began using *RH*-cDNAand poly-A cDNA from the polar night field campaign as this had repeatedly shown promising results on PCR. The three primer pairs selected for the assay were subjected to six preliminary qPCRs to optimize reaction conditions and materials. These qPCRs involved either singles or duplicates of both DNA and RNA extracts and covered all field campaigns from the time series. A primer pair targeting Dinoflagellate rhodopsin was included from the Huang et al. study in the hopes of observing in-situ evidence of Dinoflagellate-associated rhodopsin transcription (Huang et al., 2019).

qPCR was performed using a 96 well StepOnePlus™ Real-Time PCR System from Thermo Scientific™. Applied Biosystems™ SYBR™ Green PCR Master Mix was used. Reactions were set up using MicroAmp® Fast 96-Well Reaction Plates (0.1mL) from Applied Biosystems®, on an Eppendorf™ PCR Cyroblock which was otherwise kept at 0°C. Once loaded, plates were sealed using Silicone Adhesive film from VWR International™. These trials involved separate runs at 45°C and 50°C to optimize annealing temperatures.

This was followed by a temperature trial using the VeriFlex™ function on the instrument, using poly-A cDNA from the first sampling date from the field campaign (17/01/2023) as this contained the highest amounts of RNA (4ng/ μL). This resulted in unspecific amplification with many melt curves observed. Another qPCR using *RH*-cDNA was trialed for the primer pairs, opting for 56°C as this had previously displayed optimal primer activity.

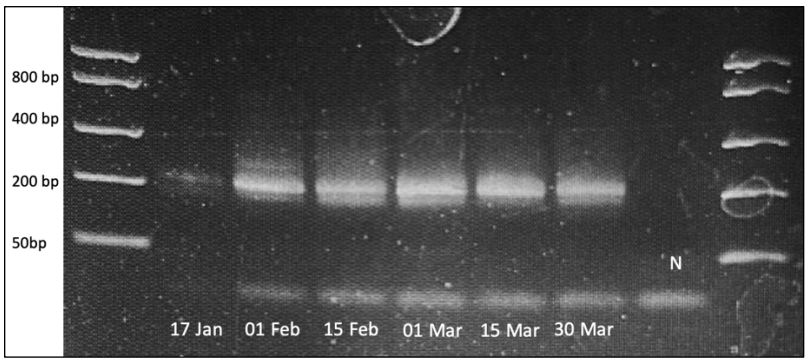


Figure 13. qPCR results for primer pair KarloF4/R4 tested on time series samples using Random Hexamer cDNA.

This deemed the best genetic material to use for the final qPCR assay (as can be seen from Figure 13) as product were of expected sizes, and uniform across the time series.

5.2 qPCR Assay

The final qPCR was set up such that each primer pair was tested with 3 replicates of *RH*-cDNA, and 3 replicates of DNA from each sampling date from the time series, and one NTC sampling date (6 per primer pair), resulting in a plate with 84 filled wells. The program used was the same as that used for all previous PCRs, save for a first 10-minute step at 95°C to activate the AmpliTaq Gold™ DNA Polymerase in the SYBR™ Green Master mix. After each qPCR run, data collected during the melt curve stage provided information on the melting temperature (T_m) of my amplicons whilst also investigating nonspecific PCR amplification (primer dimers and/or contamination) with changing template concentration. Results were validated using agarose Gel.

5.3 Standard Curve

A Standard Curve was integrated for each assay as it was evident from PCR that nucleic quantities from the environmental samples were on the detection limit. The standard curve would enable me to assess the performance of the qPCR assay by estimating its efficiency and provide limits of both detection and quantification. This required the following materials, which were used on SPRI-cleaned colony PCR products according to the manufacturer's instructions:

1. Freezer stocks of plasmid
2. Monarch® Plasmid Miniprep Kit from NEB®
3. Monarch® PCR & DNA Cleanup Mit from NEB®
4. Fastdigest HindIII from ThermoScientific™

Stock solutions of plasmids transformed with the 3 primer pairs for this assay were subject to serial dilution using 10-fold dilution steps, resulting in samples with concentrations 10^9-10^2. The standard curves were run on the same MicroAmp plates as the final assays to invalidate entropy from the reactions.

6.0 PCR and qPCR product purification

Throughout this study, PCR and qPCR products were cleaned and purified for different purposes on several occasions. Products were cleaned for nested PCR, bacterial cloning and Sanger sequencing. The method used was Solid Phase Reversible Immobilization (SPRI) using paramagnetic beads to selectively bind smaller products.
Products were prepared for Sanger sequencing such that 7.5 µl of cleaned product (by SPRI beads) was added to a 1.5 mL Eppendorf ® tube with 2.5 µl of 10µM of the forward primer used for the amplification. Products were sent to Eurofins Genomics for Sanger sequencing.

6.1 SPRI Bead cleaning for Sanger sequencing

The ratio of SPRI bead solution : PCR product was altered according to product size as early attempts at doing so with 1:1 ratio resulted in the removal of all PCR products. The following protocol was optimized during this study and is therefore being made available for further use. I optimized the protocol in the following ways: 1. Once the PCR product had been added to the SPRI bead mixture, the tubes were flicked by hand for 30 seconds then micro-centrifuged for 30 seconds to ensure thorough mixing, 2. Tube lids were closed after Ethanol (EtOH) was removed as many samples were being processed at once, therefore all lids were opened at the same time, 3. The tubes were left open for 2 minutes to ensure evaporation of residual EtOH, and 4. The tubes were left on the magnetic rack at all times when pipetting.
To test the efficacy of this protocol once optimized, it was first tested at various concentrations (0.8x, 1.0x, 1.5x, 2.0x, 2.5x) on samples containing: 2 µl of Gene Ruler Low-Range Ladder + 18 µl MilliQ-H_2O. The inclusion of this step was recommended to perfect the technique and check the viability of the products.

6.2 Gel extraction

When performing nested PCR, multiple bands of different sizes were occasionally observed on agarose gel. To determine what these represented, gel extraction was performed with subsequent cleaning of each band.

The gel extractions were performed according to the QIAquick ® Gel Extraction Kit from ©QIAGEN. Table 7 can be referred to for the primer pairs of the amplicons which were extracted. Once extracted and purified, the samples were sent for Sanger sequencing.

Table 7. Primer pairs used in Nested (first 4) and Standard PCR (last 2), and which bands were extracted (ref. Figure 14).

Primer pair	Band extracted
BlueF & OP283	Upper and Lower.
DinoGF1 & OP283	Upper and Lower.
DinoGF2 & OP283	Upper and Lower.
PhaeoF1 & OP283	Upper and Lower.
DinoGF1 & DinoGR1	Upper and Lower.
DinoGF2 & DinoGR2	Lower.

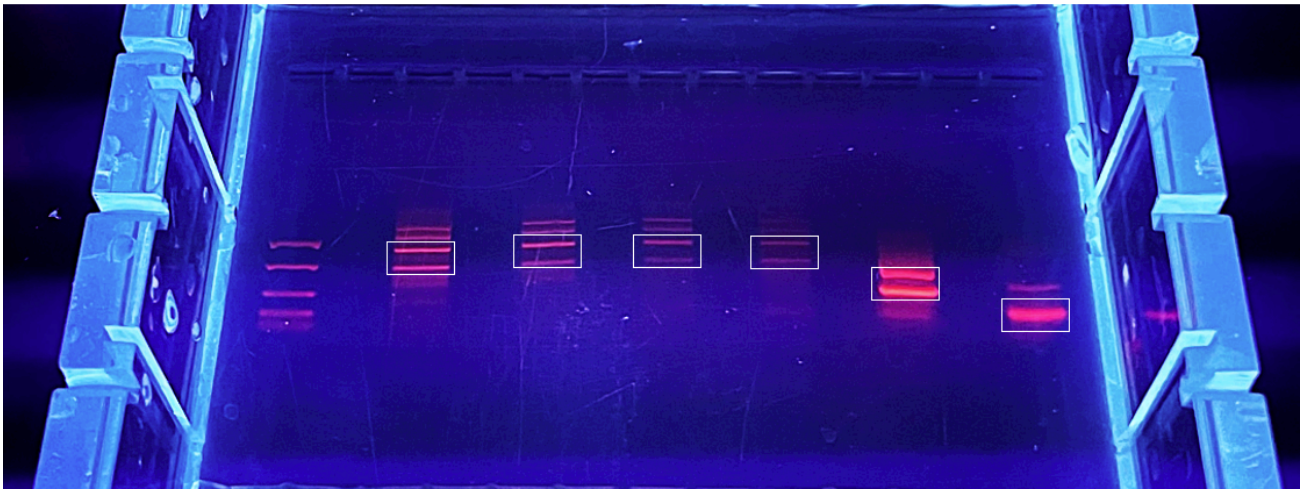


Figure 14. Ultraviolet Agarose Gel Extraction displaying multiple-sized products.

As can be seen from Figure 14, multiple sized products are observed on agarose gel, however these appear to be single products as the bands are crisp. The first 4 wells are the results from nested PCR, whereas the last 2 are from standard PCR. This is demonstrative of the specificity gained from nested PCR, as can be seen from the better separation and definition of the bands from these wells.

7.0 Data Analysis

Physicochemical environment data

All data was processed using Microsoft Excel (Version 16.75).
For ease, only data from one CTD probe (SD 208, SAIV A/S) was used for analysis. No data from the field campaign on 15th March was recorded on the SD208 probe, therefore data from the RBR*concerto*^{3} CTD was used and corrected using a linear regression model.
Data from the spectrofluorometer was plotted to visualize spectral conditions at the water surface throughout the time series. This data was later normalized using the equation below, applied to each field campaign individually. A sum of all values from each field campaign was used to normalize the data. This was done to investigate the dominance of certain light types at each date as it may not be visible when the data is plotted along the same y axis due to the difference in light intensities throughout this time.

$$Normalization = \left(\frac{Value}{Sum}\right) * 100$$

qPCR assay data

Data from the qPCR assays were manually plotted to be able to maximize data extrapolation. From the standard curve (SC), two analyses were performed on the data: 1. Amplification efficiency, and 2. Copy number of target DNA/cDNA.

The amplification efficiency (*E*) of my reactions were calculated using data from the standard curve, with the following equation: $E = -1+10^{(-1/slope)}$ for each primer pair. This was done by plotting the Cycle threshold (Ct) values on a logarithmic scale along with their corresponding serial concentrations, followed by generating a linear regression curve through the data points to calculate the slope of the trend line. The *E* value for each primer pair generated from the SC was then applied to the assay data.
Copy number of target DNA/cDNA was calculated in a two-step method. Using the linear regression models from the SC, the following equation was used to estimated copy number from the qPCR reaction:

$$Copy\ number = 10^{(Ct-Intercept)/Slope}$$

Secondly, these values were backcalculated to obtain copy number/mL filtered seawater to estimate the starting quantities in my environmental samples and to be able to compare the assays at equal measure. This was done using the following equations for DNA and cDNA, respectively:

$$DNA\ copy\ number/\ mL = \frac{qPCR\ copy\ number * (\frac{Elution\ vol.}{\mu l\ DNA\ in\ qPCR})}{Vol.\ filtered\ seawater}$$

As cDNA synthesis may interfere with correct RNA copy number estimation, the equation was altered to account for the variations in concentration which may arise with DNAse treatment and cDNA synthesis.

$$RNA\ copy\ number/\ mL = \frac{qPCR\ copy\ number * 60 * 62 * 20}{54 * 1 * 1 *\ mL\ filtered\ seawater}$$

Where 60= µL in RNA elution, 62= total µL in DNAse reaction, 20= µL in cDNA synthesis reaction, 54= µL RNA in DNAse reaction, 1= µL DNAse-treated RNA in cDNA synthesis reaction and 1= µL cDNA used in qPCR reaction.

Results

Research Aim 1: Light regime (spectral composition, intensity, and duration).

Despite the spectrofluorometer being deployed along the same vertical profile as the CTD probes, only surface data will be presented and discussed. The recordings begin on 1st February 2023 and were taken from wavelengths 379.52-872.21nm. Spectral conditions evolved drastically throughout this time series, as can be seen from Table 8.

Table 8. Spectral conditions (daylength and solar declination angle at 11:30a.m local time- CET- the average sampling time) throughout time series. Daylength obtained from https://www.timeanddate.com/, Solar declination obtained from https://gml.noaa.gov/.

Sampling date	Daylength (hours)	Solar declination (°)
17/01/2023	Twilight only.	-20.75
01/02/2023	Twilight only.	-17.11
15/02/2023	Twilight only.	-12.69
01/03/2023	07:23	-7.6
15/03/2023	11:09	-2.15
31/03/2023	15:17	4.12

As solar radiation returned to this polar environment, spectral conditions changed from scattered light, characterized as “twilight only” (defined by dim blue light), to daylength of +15 hours. Solar declination means that the sun’s rays hit the sea surface at a very oblique angle, spreading and diluting the radiation. Spectral conditions evolve accordingly, whereby spectral quality (light type) and quantity (light intensity) changed throughout the time series, as can be seen in the data from the Spectrofluorometer.

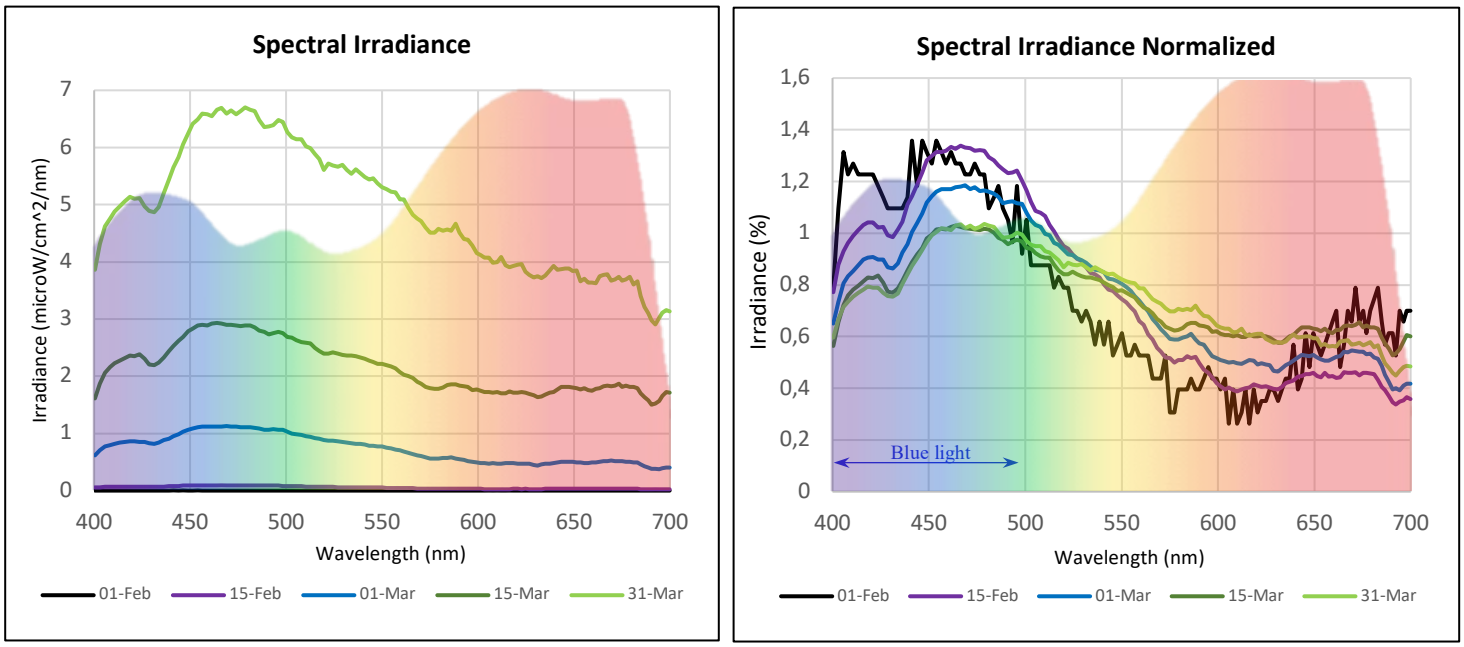


Figure 15. Spectrofluorometer data collected throughout time series at 0m. A. Spectral Irradiance values collected within PAR (400-700nm), with PAR spectral composition overlayed . B. Spectral Irradiance values normalized, distinguishing the blue-light range (400-500nm).

The results show a clear evolution of the light regime throughout the time series, demonstrated by changes in spectral composition and intensity. Spectral conditions in the water column for the first two sampling dates (1st and 15th of February) show little to no spectral irradiance, and are therefore hardly visible in comparison to the later field campaigns, with recordings indicating average values of ~0.001-0.08 µmol m^{-2} s^{-1} in the blue-light range (~400-500nm) for these dates. March 1st demonstrates a large increase in spectral irradiance, whereby values reach > 1 for the first time. Throughout the remainder of the time series, as daylength increases and sunlight returns to this environment, irradiance broadens both by spectral type and intensity, with blue-light being the dominant light-type, with a blue-light maxima of 6.7016 recorded on 31st March. When viewing the data normalized, it is further emphasized that blue light (~400-500nm) is the dominating light type throughout of the time series (Figure 15.B). The recordings from the two earliest field campaigns (1st February and 15th February) can be seen to have the highest incidences of blue-light despite the low-light conditions caused by civil twilight.
However, we can see a gradual shift to white light-dominating conditions (where all wavelengths of the visible spectrum are at equal intensity) by the last sampling date (31st March).

Research Aim 2: Design of a qPCR assay through PCR optimization

2. 1 Optimization via PCR

2.1.1 cDNA variations

In most cases, poly-A cDNA from the polar night campaign provided the most promising results when tested with rhodopsin-targeting primers. Despite obtaining multiple products of different sizes, this cDNA was used most regularly for PCR trials as *RH* cDNA often resulted in large, unspecific products from PCR (observed as a smear rather than a band) when observed on agarose gel.

Conversely, for the qPCR assay preliminary qPCRs using poly-A cDNA resulted in unspecific amplification and very large products, seen as smears and bands indicative of amplicons +800bp on agarose gel. This being a prime example of primer entropy, whereby different amplification outcomes are observed when applied to environmental samples using variants of genetic material.

2.1.2 Temperature trials

When verified on agarose gel, cleaner, temperature trials demonstrated brighter bands at higher temperatures than those observed at 50°C.

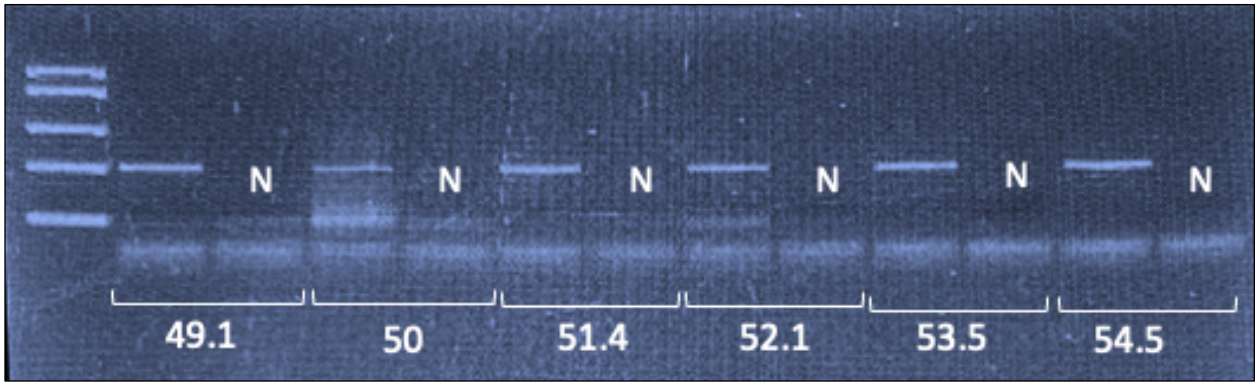


Figure 16. 1.0% Agarose gel result using 3µl GeneRuler low range ladder from Thermo Scientific™ for temperature trial for psb primer pair (photosynthesis psbA protein) for *Phaeocystis pouchetii* from 49-54.5°C, showing increasingly brighter bands with higher temperature.

The optimal temperature was used for further testing with different cDNA variants (*RH*-cDNA, poly-A cDNA) to provide clean sequences for Sanger sequencing.

2.1.3 General eukaryote and photosynthesis primers

In the first few months of the study, general eukaryote primers obtained from literature were included in almost every PCR alongside rhodopsin primers.
These often provided large, unspecific products with smeared bands on agarose gel (Figure S3. in Supplementary material) and were therefore later abandoned to focus on two high Arctic phytoplankton strains. The photosynthesis primers both designed for this study and from literature were tested on cDNA from both polar night and September at different temperatures but never provided results which were promising enough to pursue. The same was observed for the spliced-leader primer *dino-SL*, which despite being tested with many blue- and green-light rhodopsin-targeting reverse primers, no product of the expected size was seen. Instead, a low molecular weight product was evident in both sample and negative control, suggesting the formation of primer dimers.

2.1.4 Species-specific primers

The 19 species-specific primer pairs proved unsuccessful (save for one, discussed below) whereby only primer dimers in both the template and negative samples were observed when tested on poly-A cDNA from the polar night field campaign. These were therefore abandoned. The primer pair targeting the psbAK gene in *Phaeocystis pouchetti* showed promising results and underwent a temperature trial to optimize reaction parameters. However, this primer pair was later abandoned when no results were observed when tested on time series cDNA.

2.1.5 Nested PCR

Nested PCR was done using poly-A cDNA from the polar night sample made with primer OP41 (Vader et al., 1999). This was performed for all primers which: 1. Showed promising results in first round of testing (correct amplicon size + no primer dimers /product in negative), and 2. Underwent temperature trials. Forward primers chosen for the first PCR included those targeting 1. Blue-light rhodopsin, 2. General dinoflagellates, 3. Rhodopsin in uncultured eukaryotes. For demonstrative purposes, only one nested PCR shall be showed here. Primer BlueF, designed to target blue-light rhodopsin in *K. veneficum*, was paired with Rp OP283. The first PCR on poly-A cDNA from polar night resulted in very large, unspecific products, as can be seen in Figure 17.

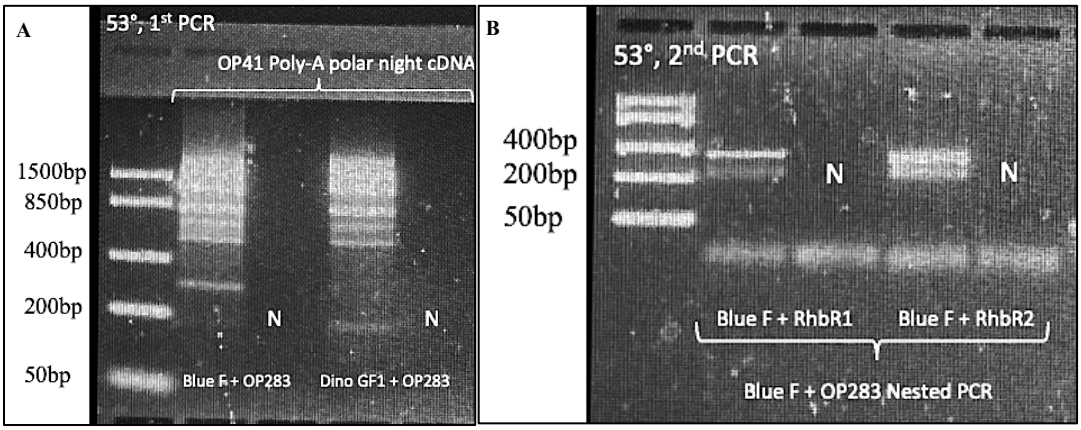


Figure 17. Results from nested PCR on 1% agarose Gel, using 3µl GeneRuler low range ladder from Thermo Scientific™. A. 1st PCR with forward primers and poly-A complementary reverse primer OP283. B. 2nd PCR with same forward primers and associated reverse primers showing smaller, more specific products.

Results from the first PCR (Fig. 17.A) can be interpreted such that the multiple bands observed are likely due to the fact that OP283 will pick up on all transcripts with a poly-A tail (of eukaryotic origin), of which the lengths of transcripts and poly-A tails present in the sample will vary greatly. As a result, this can be seen as products of varying sizes (most of which > 400 bp) . However, a single product is observed at around the ~300bp mark, as can be seen by the faint band next to GeneRuler low range ladder on the left of the figure.

When the same forward primer is then paired with its corresponding reverse primer for the second PCR (here RhbR1- blue-light rhodopsin-targeting) a product of the same size is observed. However, the second PCR rendered the band brighter, and removed the unspecific and undesired

larger amplicons from the first PCR 5 (Fig. 17.B). Results such as these served as the template for bacterial cloning after cleaning by SPRI beads due to the results clearly indicating single products.

2.1.6 Bacterial cloning

Colony PCR products sent for Sanger indicated successful hits on rhodopsin sequences for *Dinophyceae* sp., *Karlodonium veneficum, Prorocentrum donghaiense* and proteorhodpsin sequences from Uncultured marine microorganisms when using the BLAST function from NCBI. This was an unprecedented record of rhodopsin from Arctic strains of phytoplankton from PCR.

Table 9. Sanger sequencing results displaying nucleotide BLAST hits to phytoplankton rhodopsin sequences in various marine microbes.

ID	Primers	Hit	Hit ID	Query cover	Per. ident
ELY997	*BlueF/ProteoR2*	Uncultured organism proteorhodopsin	KJ926746.1	64%	67,81%
ELZ008	*BlueF/ProteoR2*		KJ929641.1	94%	75,38%
ELZ007	*BlueF/ProteoR2*	*Karlodinium veneficum* rhodopsin	MW570712.1	81%	76,14%
ELZ012	*BlueF/ProteoR2*		MW570713.1	79%	72,85%
ELZ011	*BlueF/ProteoR2*		MW570713.1	83%	75,17%
ELZ015	*DinoGF1/DinoGR1*	*Dinophyceae* sp. rhodopsin	MW570710.1	45%	95,59%
ELY998	*PhaeoF2/PhaeoR1*	*Prorocentrum donghaiense* rhodopsin	KM282617.1	95%	81,82%
ELZ013	*PhaeoF2/PhaeoR1*		KM282617.1	66%	75,18%
ELZ018	*PhaeoF2/PhaeoR1*		KY399746.1	65%	76,67%
ELZ002	*RhbF2/RhbR1*	Uncultured organism proteorhodopsin	KJ929840.1	57%	70,54%
ELZ001	*RhbF2/RhbR1*		KJ929840.1	46%	77,88%
ELZ006	*RhbF2/RhbR1*		KJ929840.1	46%	76,19%
ELZ017	*RhbF2/RhbR1*	*Karlodinium veneficum* rhodopsin	MW570712.1	63%	77,54%
ELZ016	*RhbF2/RhbR1*	Uncultured organism proteorhodopsin	KJ929840.1	50%	76,92%
ELZ020	*RhbF2/RhbR1*	*Karlodinium veneficum* rhodopsin	MW570712.1	64%	78,91%

With “Per. ident” indicating the extent to which two aligned sequences (here being those from the database and mine from Sanger sequencing results), it can be concluded that this was a strong result indicative of a resemblance to rhodopsin in *Karlodinium veneficum* and *Prorocentrum*

donghaiense from my amplicons. As the latter is not present in cold waters, this suggests that a new area of the rhodopsin gene in *Prorocentrum* sp. has been isolated, for the first time in high Arctic strains of this species.

2.1.7 Primers designed for final assay

The final set of primers designed and selected for the assay underwent a testing procedure on both *RH*-cDNA and poly-A cDNA from the polar night sample. Figure 17 shows PCR products of the three primer pairs chosen for the qPCR assay- ProroF1/R1, KarlorhodF2/R4 and KarloF4/R4- on Agarose gel when trialed with genetic material from the first field campaign (17th January). Using one replicate of *RH*-cDNA, DNA and a no template control, this was done several times to ensure reaction conditions for the assays had been optimized.

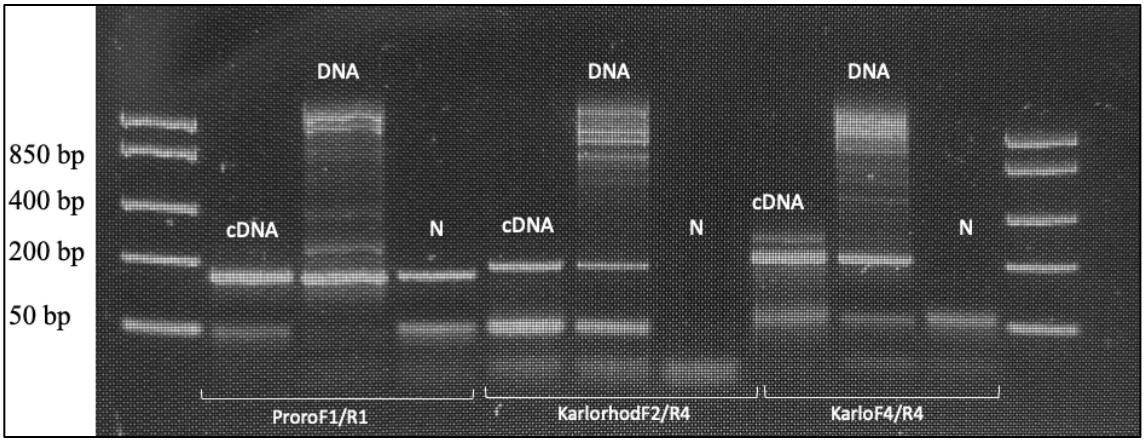


Figure 18. PCR results for primer pairs chosen for final qPCR assay displayed on 2% Agarose gel with GeneRuler low range ladder (Thermo Scientific™). Primer pairs ProroF1/R1 for *Prorocentrum* sp. and KarloF4/R4 & KarlorhodF2/R4 for *Karlodinium veneficum.*

Results demonstrate larger, less specific bands associated with DNA amplification. Product sizes from amplification from cDNA are of the expected size (~160, ~180 and ~210bp, respectively).

2.2 qPCR

2.2.1 Assay Development

Standard curves for the plasmids containing amplicons made with KarloF4/R4, KarlorhodF2/R4 and ProroF1/R1 inserts were plotted such that the amplification of the samples (recorded as ΔRn) were plotted against Cycle number. This data was plotted integrating the no template control and cycle threshold (Ct) for each experiment.

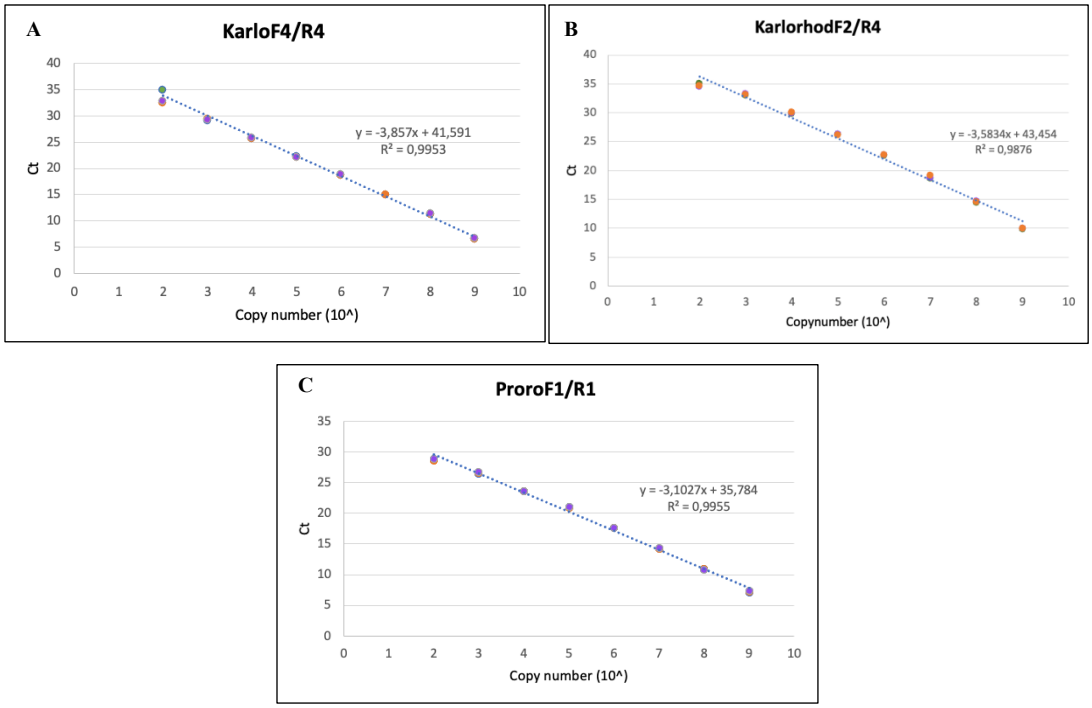


Figure 19 (A, B, C). Linear Regression equations generated from Standard Curve Cycle threshold (Ct) values. A. KarloF4/R4. B. KarlorhodF2/R4. C. ProroF1/R1

As can be seen from the results (Fig. 19), Ct levels are inversely proportional to the amount of target nucleic acid in the samples. As I have struggled with the presence of primer dimers throughout my testing, the no template control aids in understanding the baseline when interpreting amplification. Furthermore, the standard curve enabled me to calculate copy number and the amplification efficiency (*E*) of my reactions. Amplification plots for the standard curves can be found in supplementary material (Fig. S5).

Table 10. Linear regression results from Standard Curve serial dilutions plotted against Cycle Threshold (Ct) for each assay. Results are grouped as the slope, R^2, and Amplification Efficiency.

	KarloF4/R4	KarlorhodF2/R4	ProroF1/R1
Slope	-3.857	-3.5834	-3.1027
R^2	0.9953	0.9876	0.9955
Efficiency (%)	81.66	90.15	110

The amplification efficiency for the Standard Curves can be interpreted such that all primers performed sub-optimally (E ≠ 100). Desired efficiencies range from 90-110%. Anything outside of this range could be due to samples containing inhibitors (likely with environmental samples). The standard curve also confirmed that my environmental samples were in and around the detection limit.

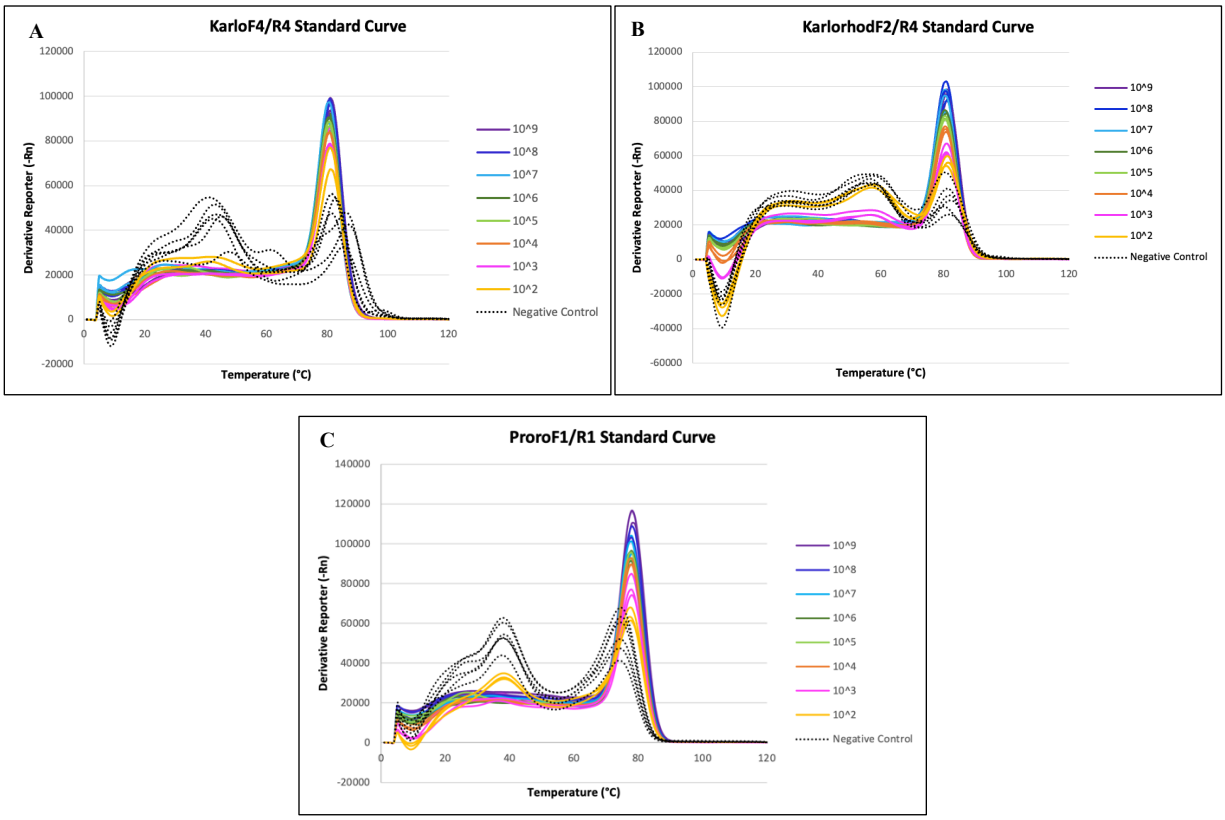


Figure 20 (A, B, C). qPCR Melt Curves for Standard Curve serial dilutions. A. KarloF4/R4. B. KarlorhodF2/R4. C. ProroF1/R1

The melt curves results enabled me to differentiate products from primer dimers. However, unfortunately, stochasticity was observed among the NTCs. Despite evidence of different degrees of fluorescence reported from the NTCs in each assay, they do indicate the presence of a similar product (primer dimer) due to their T_ms being the same.
The melt curves showed that all PCR reactions produced one specific product with a melting temperature around 80°C. In addition, primer dimers can be associated with low temperatures in both NTCs and in low concentrations of genetic material. Products in the NTC with the same melting temperature as specific products suggest contamination. To test this, a different variation of SYBR Green (POWER SYBR Green ®) and new primer aliquots was included in the KarlorhodF2/R4 assay. The use of two variations of NTC and new primer aliquots proved no difference, meaning that the primer stocks were likely contaminated. This was also confirmed by the agarose gel whereby the products in the NTC are the same size as the qPCR products (~180bp for KarlorhodF2/R4).

A clear correlation can be seen in the standard curve dilutions whereby as plasmid concentrations increase, primer dimers decrease in occurrence and the fluorescence signal which is emitted from increasingly specific products increases. The assays were successful but with a low detection limit of approximately 7 copies/mL of filtered seawater.

2.2.2 Rhodopsin gene abundance and expression

DNA and cDNA amplification are grouped per primer pair (assay) for comparison, without distinguishing dates. Amplification plots for each primer pair showing DNA and cDNA amplification separately throughout the time series can be found in supplementary material for reference (Fig. S6).

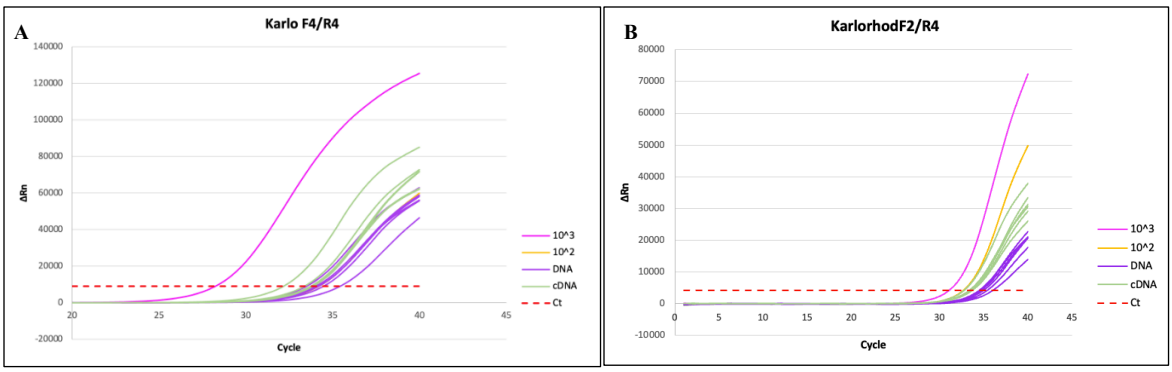


Figure 21 (A, B). qPCR Amplification plots for primer pairs targeting *Karlodinium veneficum* with standard curve serial concentrations 10^2 and 10^3 (yellow and pink curves, respectively). Cycle threshold demonstrated by dashed red line.

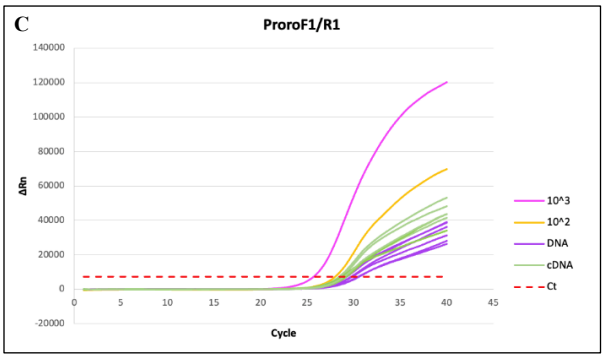


Figure 21. C. qPCR amplification plot for primer pair ProroF1/R1 targeting *Prorocentrum* sp. with standard curve serial concentrations 10^2 and 10^3 (yellow and pink curves, respectively). Cycle threshold demonstrated by dashed red line.

The amplification data obtained from the assays demonstrates a slight decrease in amplification throughout the time series, for both cDNA and DNA. Out of the two *K. veneficum* primer pairs, the assay for KarloF4/R4 demonstrated the most trust-worthy results, with slight but visible decreases in amplification throughout the time series, noticeable for both DNA and cDNA, with the 10^2 SC dilution falling in between DNA and cDNA amplification. Although not displayed here, the NTC is also well below the sample amplifications and crosses the samples' Ct later. Higher amplification in the cDNA suggests the presence of rhodopsin transcripts in the environmental samples, which decrease in quantity throughout the time series. KarlorhodF2/R4 displayed a consistent amplification relationship in both cDNA and DNA whereby a spike occurred (amplification decreased, spiked up and decreased throughout time series). However, the NTC had higher amplification than all DNA therefore rendering the

extrapolation of this amplification data as inapplicable. Furthermore, both of these primers were designed to target the same region in *K. veneficum*, but different amplification scales are observed in both cDNA and DNA (in DNA, KarloF4/R4 amplification ~5800 ΔRn, KarlorhodF2/R4 amplification ~2000 ΔRn).

The latter demonstrated a similar result to that of ProroF1/R1, whereby little genetic material in the environmental samples resulted in minimal amplification, as demonstrated by the NTC which has stronger amplification with DNA than half of the time series samples. The same aforementioned spike trend is observed in both cDNA and DNA, however amplification levels remain low overall.

When the copy numbers from the standard curve linear regression were back-calculated to obtain the environmental copy numbers (/mL of filtered seawater), copy numbers from target RNA (hereinafter referred to as cDNA) were 4 to 19 fold higher than in DNA, indicating expression in these targets (refer to Table S11 in supplementary material).

DNA copy numbers decreased throughout the time series in both *K. veneficum* assays, indicative of a successful assay given that they amplify the same target, with KarloF4/R4 being just slightly upstream of the latter (Figure S4 in supplementary showing sequence alignments). Despite the trend, DNA copy numbers in the KarlorhodF2/R4 assay appear to be roughly 2-3 fold higher than in KarloF4/R4 (i.e 17th Jan, KarloF4/R4= 17 copy numbers/mL^{-1}, KarlorhodF2/R4= 29 copy numbers/mL^{-1} and 31st Mar, KarloF4/R4= 8 copy numbers/mL^{-1}, KarlorhodF2/R4= 25 copy numbers/mL^{-1}), suggesting that the KarlorhodF2/R4 may be more specific.

The *Prorocentrum* sp. assay instead displays stochasticity in both DNA and cDNA copy numbers (no clear increase or decrease, rather fluctuation).

One trend that is observed in all three assays is a decrease in cDNA from 1st Feb. til 1st Mar., followed by a sharp increase on March 15th. After this date, both *K. veneficum* assays display a decrease in cDNA, whereas the *Prorocentrum* sp. assay demonstrates a continuing increase.

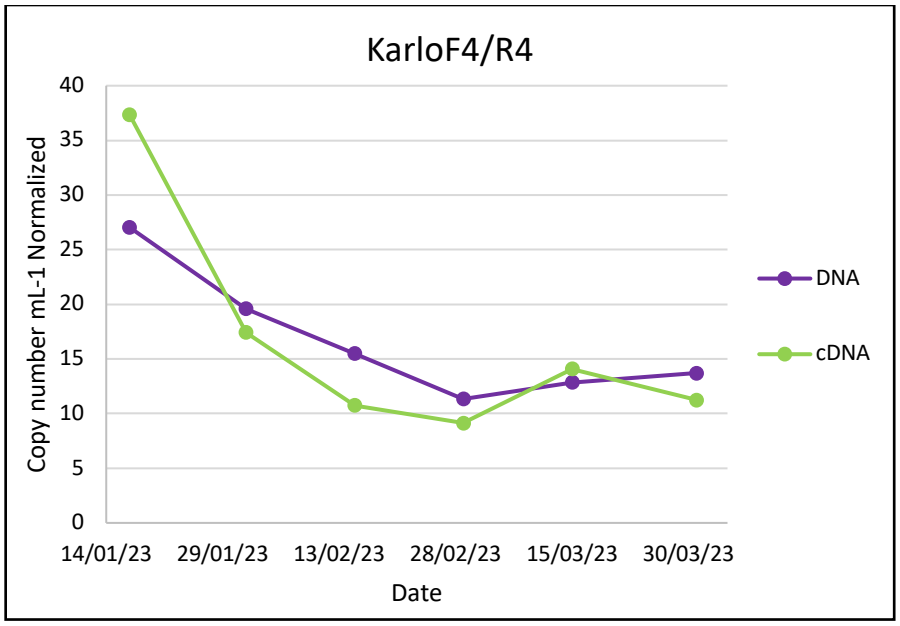


Figure 22. Normalized DNA and cDNA copy numbers per mL^{-1} filtered seawater for *Karlodinium veneficum* assay using primer pair KarloF4/R4.

When normalized using equation on page 40 (Figure S9 for assays KarlorhodF2/R4 and ProroF1/R1 in supplementary material), it is evident that the KarloF4/R4 assay seems to show the strongest correlation between DNA/ cDNA copy numbers (Fig. 22), whereby they follow the same trend.

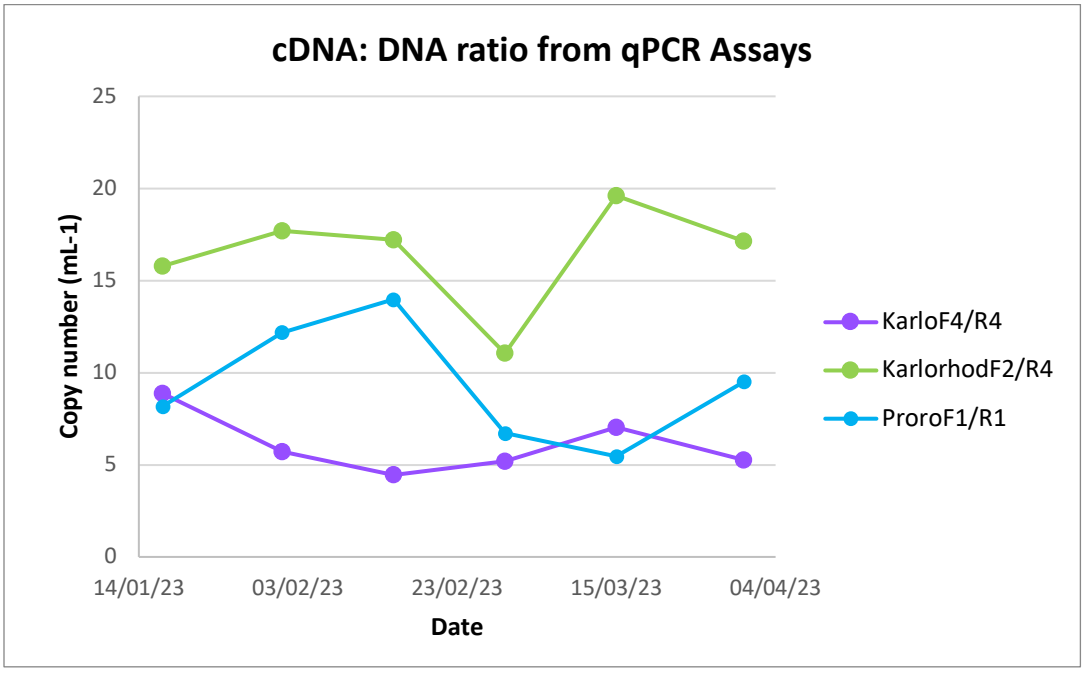


Figure 23. cDNA: DNA ratio of environmental copy numbers (mL^{-1}) for each assay.

When viewing the copy numbers per mL of filtered seawater as a ratio of cDNA : DNA (Fig. 23), it is evident that throughout the time series, rhodopsin expression fluctuates and displays different patterns in *K. veneficum* than in *Prorocentrum* sp.

When comparing the *K. veneficum* assays, difference in rhodopsin expression is observed from the beginning of the time series until 1st March, after which expression displays the same trend whereby the ratio of cDNA: DNA seems to increase, reach a high, then decrease again until 31st Mar. Meanwhile, the assay for primer pair *ProroF1/R1* resembles ratios of KarlorhodF2/R4 whereby cDNA: DNA increases til 15th February, decreases over 2 sampling dates (1st Mar & 15th Mar) before increasing again.

2.2.3 Evidence of rhodopsin gene abundance

The results from the qPCR assays were sent for Sanger sequencing, and when nucleotide and protein BLASTed on NCBI, hits were mainly observed with 1. Proteorhodopsin in uncultured bacterium, 2. Putative proteorhodopsin from either uncultured Antarctic sea ice bacterium or uncultured microorganism.

Additional hits for the primer pairs were as follows:

- ProroF1/R1: Bacteriorhodopsin-like sequences in *Sphingomonas* sp., Rhodopsin in *Synedra hyperborea*, and "Rhodopsin 1" in *Phaeocystis globosa.*
- KarloF4/R4: Rhodopsin in *Karlodinium veneficum.*

However, for each assay, hits were not concurrent between cDNA and DNA, whereby cDNA clearly outperformed DNA. Out of all three assays, KarloF4/R4 had the most hits to various rhodopsin sequences, from both cDNA and DNA when BLASTed as nucleotides and proteins. Most importantly, the KarloF4/R4 assay had a hit on *Karlodinium veneficum* rhodopsin, whereas the KarlorhodF2/R4 assay did not.

Research aim 3: Drivers of rhodopsin expression

3.1 Temperature and salinity

Only surface data from CTD probe shall be presented. Plots for 90m vertical profiles of temperature and salinity throughout the time series can be found in supplementary material (Figures S1-2).

Table 11. CTD probe data displaying depth, salinity, temperature, and fluorescence throughout time series.

Date	Depth	Salinity	Temperature (°C)	F (µg/l)
17 Jan	0,94	34,575	2,116	0
01 Feb	0,84	34,529	1,445	0
15 Feb	0,91	34,086	1,696	0
01 Mar	0,78	35,014	-0,079	0,08
15 Mar	0,84	33,960	-0,253	0,13
31 Mar	1,2894	34,72	-0,272	0,08

The highest temperature recorded was 2.2°C on 17th January, and the lowest was -0.3°C on 31st March. As can be seen in Figure S1 in supplementary material, a gradual decrease in water temperature was observed during the progression of the time series. Field campaigns on 17th January, 15th February and 31st March displayed uniform water temperatures which did not vary significantly with depth. The two remaining field campaigns (1st February and 1st March) displayed fluctuations whereby on 1st February, water temperature decreased significantly (by 0.6°C) between 10 and 50m, whereas on 1st March, water temperature dipped and spiked back at around the 50m mark. The results therefore depict an environment with decreasing water temperatures throughout the transition phase from polar night to spring.

Salinity displayed very little variation with depth throughout the time series. In fact, for all field campaigns except the last on 31st March, salinity maintained a PSU of ~34.5 from the surface down to 90m. The field campaign on 31st March displays a slightly more saline profile, with a less uniform PSU range with depth, reaching a maxima of 34.7 PSU at 90m.

3.2 Nutrients

Nutrient data was plotted such that for each nutrient, the mean of the replicates is plotted throughout the time series, with the replicates as standalone values to demonstrate the variation in the data (Fig. 24).

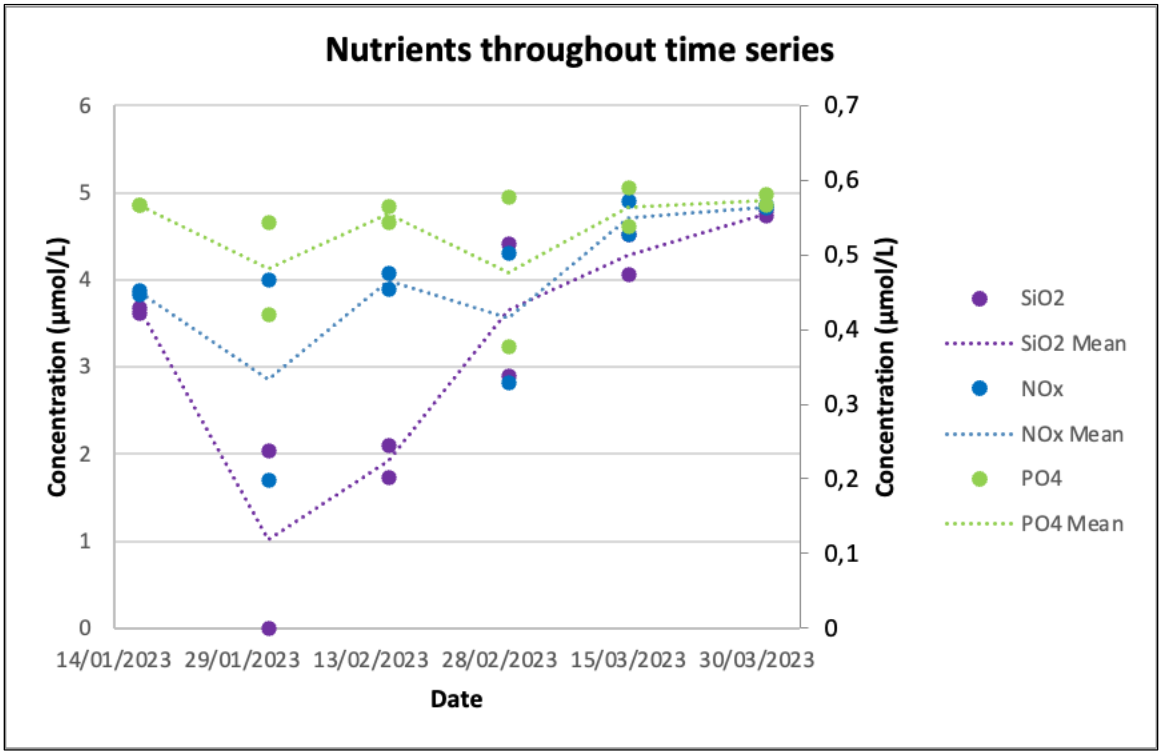


Figure 24. Nutrient analysis results for SiO_2, NO_{2+3} (NOx), and PO_4 throughout time series at 0m. PO_4 is plotted along the Secondary y-axis on the right side due to smaller values.

Nutrient concentrations for silicate and NOx displayed a gradual increase throughout the study, whereas PO_4 seems to fluctuate slightly within a very small range (0.3µmol/L variation). The results are within the expected range, and display an overall upward trend throughout the duration of the time series. The dip in all three nutrients for the first replicate from field campaign of 01/02/2023 can be interpreted either as natural variation or analysis error from the instrument. The negative value (-0.148) must be disregarded as they are considered false.

Pearson's Correlation coefficient (r) was calculated to measure the strength of the linear relationship between these nutrient abundances. NOx was strongly correlated to both SiO_2 and PO_4 with Pearson's correlation coefficients of 0.822 and 0.839, respectively. These can be

deduced as relatively strong positive correlations, meaning that nutrients are abundant in proportional ratios throughout the time series.

3.3 Fluorescence

The biological response to ambient conditions is measured by the CTD probe in terms of Fluorescence, an indirect measure of pigment concentration (Chla). Data is shown down a vertical profile of 90m to display natural fluctuation in fluorescence with depth throughout the time series.

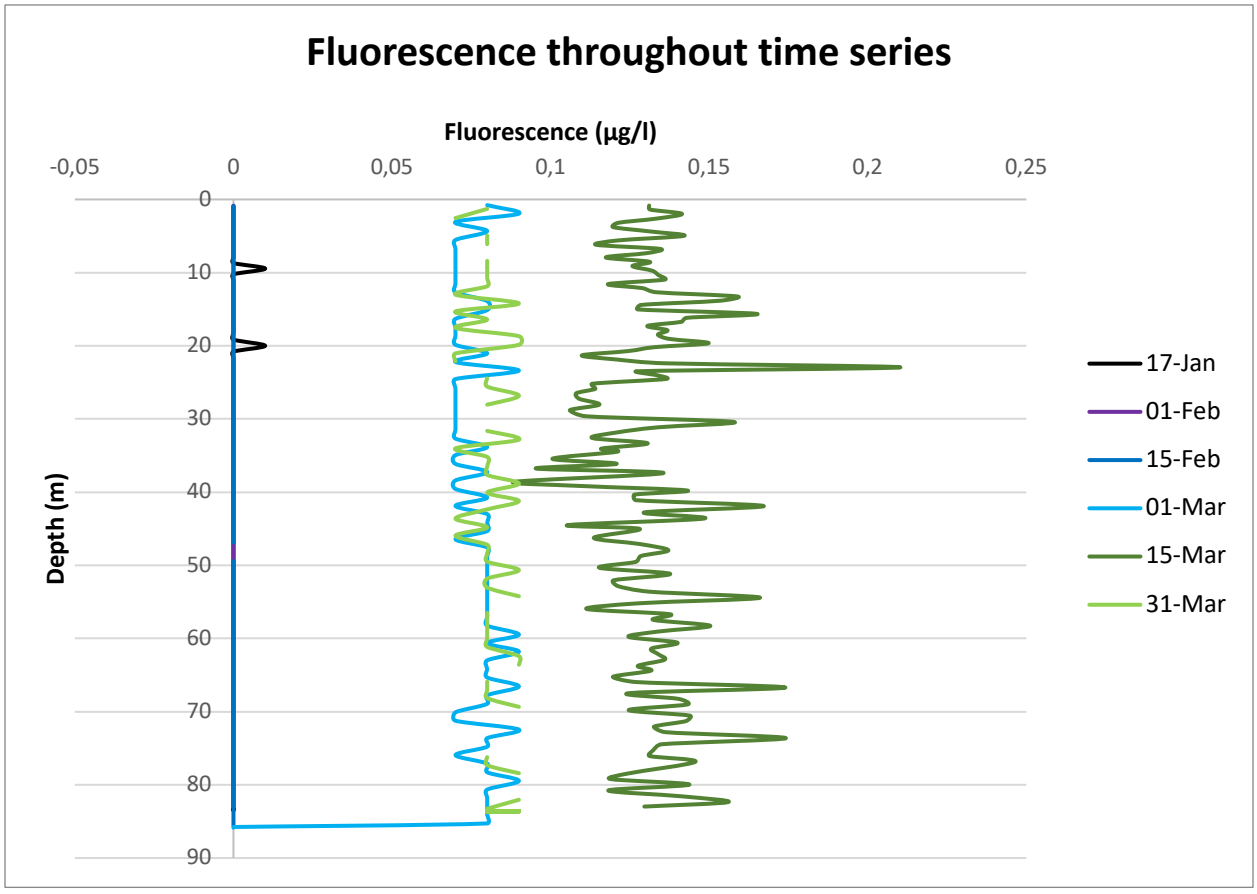


Figure 25. CTD data showing Fluorescence plotted against Depth throughout Time Series.

The data demonstrates extremely low levels of fluorescence throughout the time series, with values from the first three field campaigns being so low (almost always 0), that they are not distinguishable in Figure 26. The fluctuation and increase in fluorescence observed on March

15th is due to this data originating from the other CTD probe used during this study (RBR*concerto*^{3}) as the primary probe (from SAIV A/S) had no recordings from this date. As such, the data from the former was corrected, however a clear difference in instrument sensitivity is likely the cause for this data not fitting the trend. While the measurements indicate a slight increase in fluorescence as of March 1st (0 -> 0,08mg/L), the lack of significant difference with depth does cause for caution when extrapolating the accuracy of the data. Despite this, the increase in fluorescence can be interpreted as an increase in Chlorophyll-a throughout the time series.

Discussion

Research Aim 1: Spectral composition throughout time series

Whilst the sun remains below the horizon during polar night, solar elevation still significantly controls spectral irradiance during much of this period through atmospheric scattering of light from the sun (Connan-McGinty et al., 2022).

Very few studies have focused on underwater light climate during Arctic polar night (Berge et al., 2015; Cohen et al., 2015). However, with evidence of a significant biological response triggered by irradiance (*E*) levels corresponding to as little as 0.5% of polar night surface *E*, polar night is increasingly displaying evidence of a dynamic photoperiod (Båtnes et al., 2015). With recent research showing that Arctic zooplankton perform vertical migrations in response to changes in both the position and phase of the moon during polar night, this infers that moonlight could potentially play a major role in marine carbon sequestration during polar night (Last et al., 2016). As such, its previous perception as being a static event is now better understood as process with four different levels of light (polar twilight, civil polar night, nautical polar night and astronomical polar night), with each level defined by the sun's elevation/declination relative to the horizon (Berge et al., 2020). Each of these grades of twilight possess characteristic intensity and spectra, however they are similar in that blue wavelengths are the principle spectral composition, as is demonstrated by the data collected for this study (Fig. 15). The data demonstrates a change in both spectral quality and quantity as solar irradiance returns to a polar latitude after an extended period of darkness. Irradiance for the first three field campaigns, from January 17th- February 15th, display little to no spectral irradiance, with a sudden increase (>1) occurring on March 1st. When normalized, data from the first two field campaigns at which the spectrofluorometer was deployed (1st February and 15th February), display levels of blue light being at their highest of the whole time series, demonstrating the spectral characteristics of nautical twilight occurring at this time. As spectral composition of light is fundamental to its biological impact, the normalized data showing this trend is of great photobiological relevance. This suggests that even during the darkest photoperiods during polar night, light-harvesting activity could be sustained due to the presence of this light-type. Research has shown evidence of rapid reactivation of photosynthetic activity in communities (from samples collected at surface

level) which were exposed to light, suggesting the maintenance of (almost) fully intact photosynthetic machinery during the dark winter period (Kvernvik et al., 2018). With this in mind, it can be speculated whether or not the same observation would be made for rhodopsin, specifically in knowing that spectral conditions at this time could theoretically allow for light-harvesting activity in the blue-light range of PAR (400-500nm).

The optical properties of water are equally as important as spectral composition when discussing photobiology. Measurements of the underwater light field during polar night remain scarce due to the detection limit of commercially available sensors and the challenges posed by taking such measurements without light pollution (Berge et al., 2020). However, of the little data that is available, it is suggested that the optically clear water present during polar night is similar to that of early spring, prior to the spring bloom and extensive glacial runoff (Cohen et al., 2015). Therefore, the spectral conditions observed throughout this time series suggest great potential for light-harvesting activity due to the dominance of the blue light, traditionally believed to have the highest quantum yield of CO_2 and the most readily absorbed wavelength by photosynthetic pigments. Thus, by inferring that *K. veneficum* is adapted to low light, the Meng et al. study is of great relevance in this context as their findings show variation in rhodopsin expression in response to light spectrum and intensity (Meng et al., 2019).

Nevertheless, it is difficult to extrapolate the biological effect of spectral composition on rhodopsin expression in my target strains when little data is available on their intraspecific differences in this geographical area. However, with evidence that low blue light enhances growth rate, light absorption and photosynthetic characteristics in some phytoplankton species (Gorai et al., 2014), I do not interpret these spectral conditions as a limiting factor for light-harvesting activity (as will be discussed further below). Rather this suggests that rhodopsin expression observed from qPCR was potentially sustained throughout this time due to the dominance of blue- and green-light. Nevertheless, it leaves room for further research on the photobiological effects of spectral composition on rhodopsin expression and suggests the need for investigation of the latter in a controlled lab-environment to accurately understand these effects.

Research Aim 2: Design of a qPCR assay through PCR optimization

2.1 PCR

With the current understanding of rhodopsin expression in phytoplankton mainly attributed to studies from the Southern Ocean (Andrew et al., 2023), South and East China Sea (Ma et al., 2023; Shi et al., 2015) and Long Island Sound (H. Zhang et al., 2022), no such equivalence of understanding exists for Arctic strains of phytoplankton. Therefore, throughout the duration of this study, over 65 PCRs were performed on +100 primers designed in attempt to make a qPCR assay for rhodopsin expression in high Arctic phytoplankton

Firstly, designing primers proved challenging as they often lack specificity when used on environmental samples. Previously published PCR primers designed at a deeper taxonomic level are only "occasionally successful at discriminating target species from nontarget organisms" (Valiente Moro et al., 2009). The first stage of this study using highly degenerate primers was bound to alter the representation of taxa by amplification biases, specifically by primer binding discrimination. Despite the environmental samples being prefiltered to remove any bacterial/archaeal organisms, the eukaryotic kingdom alone displays huge discrepancy in genome sizes and ribosomal DNA copy numbers. As a result, PCR bias in environmental samples is likely to occur from multitemplate PCR, resulting in erroneous ratios between cells and PCR products (Ishii & Fukui, 2001).

Secondly, from the early stages of this study, it was evident that my environmental samples from September contained little genetic material, and the latter time series samples were on the detection limit of amplification. Despite the recent observation that Arctic microbial communities demonstrate diversity during polar night (Dąbrowska et al., 2020; Kubiszyn et al., 2017), their abundance does decrease drastically during this period. As a result, in knowing that phytoplankton is already scarce, it is unsurprising that it proved very difficult to target them with degenerate, let alone species-specific primers designed later in the study.

Similarly, species-specific primers targeting photosynthetic genes (RuBisCo cycle, psbA protein, light harvesting protein) were used in the hopes of measuring expression in rhodopsin versus photosynthetic machinery from the environmental samples. However, little to no success was obtained. The same can be said for *Karlodinium veneficum* rhodopsin-targeting primers obtained from a study. These primers showed no amplification, which can either be attributed to low *K.*

veneficum abundance at this time, or agreeing with the study from which they came, that rhodopsins display high diversity in geographically distinct patterns in the global marine environment (Huang et al., 2019).

Therefore, singling out rhodopsin transcripts from environmental samples containing as little as 1.43ng/µl of DNA whilst using primers which were: 1. Degenerate, 2. Of non-Arctic origin, or 3. Targeting species whose abundance we are unsure of leaves much room for improvement in further research.

Whether or not it was due to the minimal amount of genetic material from the environmental samples or sub-optimal primer design due to limited sequence availability, many of the primers contributing to the design of this assay displayed stochastic patterns which proved difficult to troubleshoot.

Nevertheless, this did allow for optimization of primer design and PCR (primer concentration/ ratio, temperature trials, genetic material and lack of primer dimer formation), which allowed for an encoding region to be further targeted through the design of non-degenerate, region-specific primers for eventual use in the qPCR assay.

2.2 qPCR Assay

This Assay shall be discussed using MIQE Guidelines (*M*inimum *I*nformation for Publication of *Q*uantitative Real-Time PCR *E*xperiments) in terms of efficiency, specificity and sensitivity and the standard curve (Bustin et al., 2009).

2.2.1 Efficiency

PCR/qPCR efficiency depends on many factors. As has been established, difficulty in designing primers for environmental samples and low genetic material are likely to have caused this result. Additionally, aside from ProroF1/R1, the other two primers are slightly degenerate, promoting variation in melting temperatures, therefore affecting the kinetics of the qPCR and likely the efficiency, a result even more likely to occur when used on environmental samples.

The efficiency of these assays was determined using a linear regression from the standard curve, the relationship of which is expected to have a negative slope (slope of -3.3 demonstrates a perfect doubling of the number of DNA template molecules in each step of the PCR, as can be seen in standard curve plots (Fig. S5) in supplementary material. A properly designed assay shall, in the absence of interfering substances in the sample matrix, amplify target DNA with at least 90% efficiency (Svec et al., 2015). As the standard curves were run on the same MicroAmp plates as the assays to avoid stochasticity, we apply their efficiency to the assays.
As such, all primers used in this assay performed just outside the optimum range for qPCR amplification (KarloF4/R4= 81.66%, KarlorhodF2/R4= 90.15% and ProroF1/R1= 110%). Despite sub-optimal efficiency, each individual assay displayed consistent patterns (i.e single product, same T_m, more cDNA amplification than DNA), which suggests that the assay could be validated as quantitative with more material from environmental samples. The amplification observed in the negatives is homologous to that in the 10^2 concentration from the standard curve, indicating either contamination or entropy associated with primer dimers. However, the extrapolation of the efficiency from this assay must be done with caution as previous studies have shown that efficiency varies with the PCR/qPCR template and can show diversity between standard curve and environmental samples (Brankatschk et al., 2012).

2.2.2 Specificity and Sensitivity

Achieving both primer specificity and sensitivity in this assay proved to be an ambitious feat given the time constraints and lack of rhodopsin sequence availability which could be applied to this geographic area.
The specificity of these primers has already been briefly argued due to two of them being slightly degenerate, promoting the possibility of primer mismatch. Aditionally, a comparison between the success of the assays (in terms of specificity and sensitivity) between the two species should be done with caution due to the fact that rhodopsin in *K. veneficum* and *Prorocentrum* sp. is believed to have different origins (Meng et al., 2019). As such, applying the same methods to both species could favor one and not the other.
The target specificity of any qPCR assay is determined by the design of the primers allowing quantification of taxonomic gene markers present within a mixed community from the domain

level down to the quantification of individual species (Smith & Osborn, 2009). The Sanger sequencing results revealing both specific and unspecific hits to rhodopsin in several phytoplankton strains and uncultured bacterium can be interpreted in several ways. Firstly, it is likely that my target is not in the database. Due to the current lack of sequences and research on rhodopsin in high arctic phytoplankton, this is very likely. Therefore, the hit results are sequences with which my amplicons share the highest sequence similarity. Secondly, this can be interpreted as supporting evidence that rhodopsin may have bacterial origins supported by the theory of horizontal gene transfer (Shi et al., 2015), and that the sequence is somewhat conserved across phytoplankton phyla. Thirdly, despite prefiltering, suspended organics can interfere with DNA extraction and reduce recovery of target molecular markers. Additionally, the same sized products observed in the no template controls as in the samples suggest contamination. Going forward, the inclusion of a Taq-Man probe would greatly improve the specificity of this assay, however this would require currently unavailable knowledge on the specific location (and whether it is conserved) of rhodopsin in high arctic phytoplankton strains.

Sensitivity in these assays is overshadowed by the effect of the SYBR Green Mastermix. This dye emits light only when bound to double-stranded DNA, but hence allows detection of DNA in a sequence-independent way. This results in both specific and nonspecific PCR fragments being measured as fluorescence. Due to the stochastic tendencies of the primers and their affinity for primer dimers, the sensitivity of SYBR Green is likely the cause for the separate distinct melting peaks (T_ms of ~40-50 and ~80°C) observed in the melt curves. Therefore, recreating the assay using a Mastermix other than SYBR Green in combination with a Taq-Man probe would likely increase the specificity and sensitivity of the assay.

Additionally, sensitivity and accuracy of PCR/qPCR using SYBR Green is affected by cDNA synthesis conditions (as was observed through different results obtained when using poly-A or *RH* cDNA during preliminary PCR trials) (Lekanne Deprez et al., 2002). Going forward, it would be beneficial to utilize the method from the aforementioned study and experiment with DTT concentration during cDNA synthesis, diluted vs. undiluted cDNA and testing different buffers to see if primer sensitivity could be enhanced.

2.2.3 Limitations of the Standard Curve

The standard curve (SC) is a "mathematical cornerstone for estimating concentrations of target genes from fluorescence data measured in PCR" (Schmidt et al., 2023). However, this method assumes that the efficiency is the same for both the standard and the sample target template, whereas in most cases, the sample template is different from the standard template used to prepare the standard curve. The standards originate from purified plasmids, whilst my samples are environmental, containing many different species and potential inhibitors. Thus, this introduces the possibility of increased quantification errors.

Simply mapping Ct values to point-estimates of concentration can be accurate. However, this method falls short when doing so with values indicative of low concentrations, especially when considering that this assay employed degenerate primers. Bru et al., 2008 observed that primer mismatches of a single base decrease *E* can cause underestimation of the actual copy numbers by a factor of 1000 (Bru et al., 2008). As such, the low *E* observed from these assays and the knowledge that my environmental samples contain little genetic material render the validity of the SC debatable when used for degenerate primers.

In comparison to the SC, my samples likely display evidence of the "Monte Carlo" effect. This describes the inherent limitations of PCR amplification from small amounts of any complex template. This effect is created by small and random differences in amplification efficiency between individual templates, a likely occurrence in environmental samples. Its principles are that the lower the abundance of any template, the less likely its true abundance will be reflected in the amplified library (Karrer et al., 1995). Nevertheless, the SC confirms that primers have worked during plasmid transformation and that BLAST results from NCBI are accurate. Going forward, a solution to the Monte-Carlo effect in environmental samples is to introduce the "One-Point calibration" (OPC) method as an alternative for absolute quantification from qPCR. This method accounts for template-related variability of *E* by correcting for differences in *E* between sample and standard (Brankatschk et al., 2012).

Despite prefiltering, suspended organics can interfere with DNA extraction and reduce recovery of target molecular markers. Additionally, the same sized products were observed in the no template controls as in the samples, likely due to contamination.

Primers copied from a study targeting rhodopsin in *K. veneficum* strains from China (Meng et al., 2019) also proved unsuccessful in this assay. This can be explained such that strains are distinctive of geographic region, and possibly that rhodopsin displays distinct expression patterns on the geographical level.

Research Aim 3: Drivers of rhodopsin expression

The amplification results of all three assays show that target copy numbers in my samples are either equal to or below that of the lowest concentration (10^2) of the standard curve (Figure 21 A, B, C). As such, we cannot fully trust these results as below this threshold, the assay is no longer quantitative. However, I observed consistency with the primer pairs whereby amplification was always higher for cDNA than in DNA, agarose gel validated the amplification of single products, and Sanger sequencing confirmed hits on *Karlodinium veneficum* rhodopsin. This suggests that the assay shows promise, but simply requires higher quantities of genetic material.
When discussing the results in terms of copy numbers, the consistent magnitude of difference observed between DNA and cDNA copy numbers (minimum difference= 4 fold, maximum difference= 19 fold), suggests that rhodopsin expression is occurring and fluctuating throughout the time series. However, the DNA copy numbers/mL filtered seawater suggest very low abundance of both *Karlodinium veneficum* and *Prorocentrum* sp. (7-17, 17-29 and 7-18 for KarloF4/R4, KarlorhodF2/R4 and ProroF1/R1, respectively, from minimum to maximum copy numbers recorded). Given that KarloF4/R4 and KarlorhodF2/R4 amplify the same target, the differential copy numbers of DNA observed between them (sometimes 2-3-fold higher in KarlorhodF2/R4) renders the specificity of these assays up for debate. Additionally, when referring to Figure 24 showing the ratios of cDNA to DNA copy numbers/mL in all three assays, it is evident that those for *K. veneficum* do not show the same trends. As such, it is difficult to draw conclusions with regards to the relationship between rhodopsin gene abundance and expression for *K. veneficum*. However, as both assays do demonstrate an overall gradual decrease in DNA copy numbers, this could be interpreted as an overall decrease (although very insignificant in terms of total biomass) in *Karlodinium veneficum,* which is concurrent with the pre-bloom environmental conditions observed favoring diatom growth and proliferation. Furthermore, the KalorhodF2/R4 assay only displayed hits (NCBI) on cDNA and not DNA.

BLAST results demonstrated 81.25 and 52% similarity with a putative proteorhodopsin gene in both uncultured microorganism and bacterium for nucleotides and proteins, respectively (Accession numbers KJ937502.1 and AJD16814.1). In comparison, KarloF4/R4 demonstrated the most relevant hit, exclusively on a protein BLAST from DNA Sanger sequencing results, on *Karlodinium veneficum* rhodopsin with 87.50% similarity (Accession WBQ85996.1 and WBQ85997.1). As such, this assay is the most trustworthy between the two for rhodopsin in dinoflagellate *K. veneficum* and is the first to record evidence of rhodopsin occurrence in a high Arctic strain of phytoplankton.
The performance of primer pair KarlorhodF2/R4 can therefore be interpreted as lacking specificity for application in strains of high Arctic phytoplankton, despite having clearly amplified unspecific RNA rhodopsin transcripts from the complex matrix within the environmental samples. This also accounts for the drastic difference in copy numbers observed between the two assays for this strain.

The assay for *Prorocentrum* sp. demonstrated large differences in cDNA to DNA ratios, however overall the copy numbers/ mL of DNA were indicative of very low abundance. The only trend that can be suggested is for the 3 first field campaigns (spanning January 17th-February 15th) where a clear increase in cDNA: DNA occurs, suggesting up-regulation of rhodopsin expression in *Prorocentrum* sp. cDNA demonstrated Hits from a protein BLAST on rhodopsin in *Phaeocystis globosa* (71% per. similarity) and *Prorocentrum donghaiense* (66%) (Accession numbers AEP68177.1 and ASA40332.1), this being the first ever record of rhodopsin in these species for strains of Arctic origin. *Phaeocystis pouchetii* is a key contributor to the spring bloom in Svalbard (Vader et al., 2015), but *P. globosa* is characteristic of more temperate waters. Therefore, this could be indicative of a first ever record of rhodopsin in high Arctic strains of *P. pouchetii*, who possess sequence similarity with temperate strains of the same genus. However, it is difficult to infer a relationship between rhodopsin gene abundance from the DNA or expression from the cDNA in *Prorocentrum* sp. (and potentially *Phaeocystis pouchetii*) given the very little copy numbers from DNA. Nevertheless, disentangling the environmental conditions throughout this time series enables me to investigate them as drivers of rhodopsin expression in relation to the trends observed.

Primary productivity here is limited by two factors which do not constrain this process simultaneously (Popova et al., 2010). The first, nutrient availability, (mainly dissolved nitrate) is the primary limiting nutrient for phytoplankton in the Arctic Ocean (Lewis et al., 2019) due to its role in chlorophyll synthesis. It imposes its limitation due to the rapid uptake and in turn depletion by diatoms of accumulated dissolved nitrate originating from winter mixing in the early onset of the spring bloom. The second, short-wave radiation at the ocean surface, imposes its limitation on the timing and magnitude of the spring bloom as ambient light preceding this time is insufficient to support photosynthesis.

Nutrient concentrations highly influence community composition, specifically in the poles where environmental fluctuations are extreme on a seasonal scale (Wietz et al., 2021). However there are only a few studies available which have investigated nutrient budgets in the Arctic (Torres-Valdés et al., 2013). Factors that regulate and limit expression in retinal-based photosystems in phytoplankton remain poorly studied, with the majority of research originating from bacterial PPRs. Nevertheless, given the theory that microbial rhodopsins arose from horizontal gene transfer from bacteria, the findings from these studies remain relevant for application in eukaryotes. For example, it was reported in *Vibrio* strain AND4 that rhodopsin gene expression was affected by nutrient limitation whereby in nutrient-limited media, rhodopsin gene expression was strongly up-regulated leading to increased survival (Akram et al., 2013). Similar studies have been conducted to investigate the effect of rhodopsins on population growth (Gómez-Consarnau et al., 2010; Lami et al., 2009; Steindler et al., 2011) and indicated that at least in some bacterial strains, rhodopsin enhances survival under nutrient deficiency (Shi et al., 2015). Nevertheless, research is emerging on the topic whereby Marchetti et al. 2015 suggest that rhodopsin-based phototrophy could account for a proportion of energy synthesis in marine eukaryotic photoautotrophs when photosynthesis is compromised by low iron availability (Marchetti et al., 2015).

Therefore, from the limited data available, I can deduct that my results fall within the ‘expected range’ for the time of year given the aforementioned ambient conditions (as confirmed by IsA data). The IsA samples were collected at 25m, a factor which I do not assume to affect my comparison due to the hydrographic uniformity in the upper water column during polar night. The data from the IsA time series shows natural variations in nutrient concentrations throughout

the spring field campaigns (2012, 2013, 2014). Despite occasional variation on a small scale, my results generally fall within the same range as IsA data (nitrate= 5.41-6.65µmol/L, silicate 2.3-3.48 µmol/, and phosphate 0.57-0.6 µmol/L) whereby throughout the time series, NOx= ~2.8-4.9 µmol/L, silicate= ~1.8-4.8 µmol/L and phosphate= ~0.4-0.6 µmol/L, demonstrating that nutrient concentrations always show natural variability, especially on an annual basis. Despite small fluctuations, silicate and nitrate/nitrite do display an overall increase in concentration throughout the time series. When using the IsA data as a reference, I can deduct that my values are indicative of a well-mixed pre-bloom water column, the result of thorough nutrient accumulation over winter.

The lack of significant fluctuation in nutrient concentrations in comparison to cDNA:DNA ratios suggest that nutrients to do not appear to be a limiting factor on rhodopsin abundance or expression (partially due to the present lack of research and understanding on how this would manifest itself). Nevertheless, this study leaves room for speculation on whether the occurrence of rhodopsin in phytoplankton could promote enhanced survival strategies in nutrient-limited conditions, as it does in bacteria.

With this time series being conducted whereby an Arctic location is coming out of polar night, spectral composition is no longer a limiting factor, as is demonstrated in the data by a gradual shift from blue to white light, and the previously discussed dominance of blue light in the early field campaigns. However, the effect of light regime, more specifically the sudden onset of higher light intensities does display a concomitant decrease in rhodopsin expression for all three assays on March 1st, which can be associated with the significant increase in spectral irradiance which occurred on this day. Phytoplankton regulate internal pigment concentrations in response to nutrient and light availability (Graff et al., 2016). During conditions of low-light, phytoplankton maximize light absorption and photosynthetic capacity by increasing the size of the photosynthetic unit (PSU) (Perry et al., 1981), whilst also increasing pigment concentration (Moore et al., 2006), amongst many others (Wacker et al., 2015). Though this is known for photosynthetic machinery, the study Shi et al., 2015 remains the only one of its kind for retinal-based photosystems, which suggests that rhodopsin may serve as an adaptation to low-light in *Prorocentrum donghaiense*. Their findings that rhodopsin-promoting light conditions were similar to that of turbid marine habitats during a spring bloom, suggest that this gene may function to compensate for light-limited

photosynthesis in this dim environment (Shi et al., 2015). Therefore, in accordance with this theory, my data could be indicative of higher rhodopsin transcription during the periods of low spectral irradiance (17th January-1st March), which is then negatively affected by the sharp increase in irradiance observed as of the latter date. This leaves room for exploring this theory in high Arctic eukaryotic phytoplankton strains as the twilight phases associated with polar night could favor retinal- over plastid-based photosystems, which would be light-limited during these times.

Specifically in polar habitats, solar radiation can increase so much during this period of spectral transition that phytoplankton must protect themselves against photodamage (Leu et al., 2007; Sakshaug, 2004). The data from the assays demonstrates a sharp decrease in cDNA:DNA ratios on March 1st coinciding with spectral recordings from the spectrofluorometer which indicate the first recordings of >1 on this date. To cope with instantaneous light stress, phytoplankton deploy diverse mechanisms involving short- and long-term physiological changes (Kvernvik et al., 2020). The immediate responses are manifested as increases in nonphotochemical quenching of excitation energy, resulting in an overall decrease in stress derived from electron pressure and reactive oxygen species (Lavaud & Goss, 2014). Over longer time scales, a shift occurs whereby photosynthetic pigments are replaced with photoprotective pigments to minimize photoinhibition (Han, 2000). However, large differences in acclimation capacities towards high-light has been observed across phytoplankton, indicating species-specific mechanisms for coping with environmental stressors (Kvernvik et al., 2020). The decrease in cDNA: DNA observed on March 1st could therefore be interpreted as downregulation of rhodopsin expression in response to the sudden high rates of spectral irradiance occurring at the sample site, similar to photoprotective patterns usually deployed in phytoplankton for photosynthetic machinery. As this decrease is observed for both *K. veneficum* and *Prorocentrum* sp. assays, this could be indicative of similar photoadaptation mechanisms by both species. Additionally, the trends in rhodopsin expression in assay KarloF4/R4 for *Karlodinium veneficum* throughout the time series (seen as a decrease in copy numbers/mL filtered seawater) can be correlated to the findings and novel concept proposed by Meng et al., 2019. Firstly, the sharp decrease in cDNA copy numbers/mL on March 1st agree with their findings that this gene is significantly inhibited under high light. Secondly, the overall decrease in rhodopsin expression observed can be interpreted as supporting evidence for their

proposed novel concept that rhodopsin can act as an alternate energy source during conditions of dim light when plastid-based photosystems are otherwise light-limited (Meng et al., 2019). Additionally, these findings correlate with those from the associated study by Wutkowska et al. 2023 study which reported overexpression of rhodopsin-associated genes during polar night.

Salinity shows a very small fluctuation throughout the time series (34.3-34.7 PSU), as would be expected as fluctuations in this area are controlled by ice and brine production in winter, and subglacial discharge in Spring (Nilsen et al., 2008; Vonnahme et al., 2021). Given that sea-ice does not form here (the IsA sampling station), these results demonstrate that wintertime conditions in the upper water column in coastal waters is characterized by rather uniform levels of salinity (down to 90m). Therefore, it is demonstrative of the expected ambient environmental conditions preceding the spring bloom in this area and as such the lack of fluctuation in salinity is therefore nullified as being a driver of rhodopsin gene abundance and expression during this time series.

Coastal and fjord waters around Svalbard reach their coldest temperatures around February/March, the period when seasonal fast-ice forms (Wang et al., 2015). Even though Isfjorden does not form sea-ice in winter, the data demonstrates ambient environmental conditions expected in winter for a polar latitude- increasingly cold-water temperatures. Specifically, the data shows a drop in water temperature of 2°C from the first to the last field campaign. However, the subsurface water temperature recorded on March 1st does display a drastic decrease compared to the previous field campaigns whereby it was -0.079°C at 0.78m, compared to 1.696°C two weeks prior. Temperature is a basic property of the environment that directly affects photosynthetic processes as well as the ability of algae to adjust their photosynthetic apparatus to changes in resource availability (Markager et al., 1999). In general, low temperatures can modify cellular metabolism and in turn cause a decrease in enzymatic function, whilst also changing the nutrient requirements for growth in phytoplankton (Anderson et al., 2022). As such, it can be hypothesized that the decreases in cDNA: DNA ratio on March 1st as well as the decrease in subsurface water temperature could be indicative of rhodopsin expression patterns concomitant to those observed in photosynthetic machinery whereby low temperatures result in light-limited photosynthetic rates to become temperature dependent (Tilzer et al., 1986). However, as research is lacking on the temperature-dependent dynamics of retinal-

based photosystems in eukaryotic phytoplankton, we can only speculate the association between this environmental factor and inherent rhodopsin expression patterns. Similar to the suspected photobiological effects of dim light, this leaves room for further investigation in a controlled lab environment to test the relationship between temperature and retinal-based photosystem expression.

Future Studies

Should this research be continued, there are several ways in which the methodology could be improved to maximize results. Given that much remains unknown on this topic, many areas of research should be integrated to further our understanding on the light-harvesting capabilities of rhodopsin in a light-limited environment.

Starting from the field campaigns, collecting larger quantities of seawater (> 20L) would provide more material as phytoplankton are known to be scarcely abundant at this time of year, as was seen with samples being on the detection limit for amplification. Additionally, performing cell counts/ biomass measurements throughout the time series would provide invaluable knowledge on phytoplankton abundance which could later be attributed to target DNA copy numbers and easily quantified in comparison to rhodopsin expression (cDNA copy numbers).

Performing metabarcoding on all samples collected throughout the time series would provide information on community composition, Chlorophyll-a, and nutrients, enabling the study of possible limiting environmental factors in rhodopsin abundance and expression. Additionally, Algal cultures for which species-specific primers and Taq-Man probes could be designed would remove the caveat that is sample unspecificity in environmental samples.

Finally, the integration of photosynthesis-targeting primers either from literature or designed to target specific high Arctic phytoplankton strains as such primers could provide insight on the differences in expression patterns (if any) between retinal- and plastid-based photosystems, and how these may be quantitatively compared in different light regimes. Finally, data interpretation must take into account the time constraints, lack of sequence availability and difficulty in designing primers for environmental samples of Arctic origin which are known to have low phytoplankton abundance at this time.

Conclusion

This study aimed to identify marked spectral contrasts through dynamic variability in rhodopsin activity. The use of degenerate primers in qPCR demonstrates how the complexity and composition of environmental samples can challenge accurate quantification by qPCR. However, the observed consistency in higher amplification from cDNA does suggest the potential of this assay, whilst emphasizing the need to test it on environmental samples containing higher amounts of genetic material. Despite this, the spectral conditions observed throughout this time series suggest no reason for photobiology from rhodopsin to cease. Contrarily, they support the idea that the dominance of blue-light from the several phases of twilight during this time series could in fact sustain light-harvesting activity associated with retinal-based photosystems in phytoplankton. Nevertheless, fluctuations in environmental conditions, namely sudden changes in subsurface water temperature and light intensity, can be hypothesized as drivers which downregulate rhodopsin expression in the two high Arctic strains of phytoplankton which were investigated. When comparing these proteins to their photosynthetic counterparts which utilize 30 plastid enzymes forming a complex harvesting system, it is evident that these structurally simple light-harvesting proteins warrant additional research to further understand energy acquisition in marine habitats, specifically in light-limited but highly productive areas such as the Arctic Ocean. As this is the first study to design a qPCR assay for high Arctic phytoplankton, I hope to provide an invaluable contribution to this field and shed a light on the potential of non-plastid-based photosystems as sources of biogenic carbon during a dynamic spectral period previously assumed to be void of life.

References

Akram, N., Palovaara, J., Forsberg, J., Lindh, M. V., Milton, D. L., Luo, H., González, J. M., & Pinhassi, J. (2013). Regulation of proteorhodopsin gene expression by nutrient limitation in the marine bacterium *V ibrio* sp. AND 4. *Environmental Microbiology*, *15*(5), 1400–1415. https://doi.org/10.1111/1462-2920.12085

Anderson, S. I., Franzè, G., Kling, J. D., Wilburn, P., Kremer, C. T., Menden-Deuer, S., Litchman, E., Hutchins, D. A., & Rynearson, T. A. (2022). The interactive effects of temperature and nutrients on a spring phytoplankton community. *Limnology and Oceanography*, *67*(3), 634–645. https://doi.org/10.1002/lno.12023

Andrew, S. M., Moreno, C. M., Plumb, K., Hassanzadeh, B., Gomez-Consarnau, L., Smith, S. N., Schofield, O., Yoshizawa, S., Fujiwara, T., Sunda, W. G., Hopkinson, B. M., Septer, A. N., & Marchetti, A. (2023). Widespread use of proton-pumping rhodopsin in Antarctic phytoplankton. *Proceedings of the National Academy of Sciences*, *120*(39), e2307638120. https://doi.org/10.1073/pnas.2307638120

Båtnes, A. S., Miljeteig, C., Berge, J., Greenacre, M., & Johnsen, G. (2015). Quantifying the light sensitivity of Calanus spp. during the polar night: Potential for orchestrated migrations conducted by ambient light from the sun, moon, or aurora borealis? *Polar Biology*, *38*(1), 51–65. https://doi.org/10.1007/s00300-013-1415-4

Béjà, O., Aravind, L., Koonin, E. V., Suzuki, M. T., Hadd, A., Nguyen, L. P., Jovanovich, S. B., Gates, C. M., Feldman, R. A., Spudich, J. L., Spudich, E. N., & DeLong, E. F. (2000). Bacterial Rhodopsin: Evidence for a New Type of Phototrophy in the Sea. *Science*, *289*(5486), 1902–1906. https://doi.org/10.1126/science.289.5486.1902

Bercel, T. L., & Kranz, S. A. (2022). Effects of spectral light quality on the growth, productivity, and elemental ratios in differently pigmented marine phytoplankton species. *Journal of Applied Phycology*, *34*(1), 185–202. https://doi.org/10.1007/s10811-021-02653-3

Berge, J., Johnsen, G., & Cohen, J. H. (Eds.). (2020). *POLAR NIGHT Marine Ecology: Life and Light in the Dead of Night* (Vol. 4). Springer International Publishing. https://doi.org/10.1007/978-3-030-33208-2

Berge, J., Renaud, P. E., Darnis, G., Cottier, F., Last, K., Gabrielsen, T. M., Johnsen, G., Seuthe, L., Weslawski, J. M., Leu, E., Moline, M., Nahrgang, J., Søreide, J. E., Varpe, Ø., Lønne, O. J., Daase, M., & Falk-Petersen, S. (2015). In the dark: A review of ecosystem processes during the Arctic polar night. *Progress in Oceanography*, *139*, 258–271. https://doi.org/10.1016/j.pocean.2015.08.005

Brankatschk, R., Bodenhausen, N., Zeyer, J., & Bürgmann, H. (2012). Simple Absolute Quantification Method Correcting for Quantitative PCR Efficiency Variations for Microbial Community Samples. *Applied and Environmental Microbiology*, *78*(12), 4481–4489. https://doi.org/10.1128/AEM.07878-11

Brey, T. (2011). E. Sakshaug, G. Johnsen, and K. Kovacs (eds): Ecosystem Barents Sea: Tapir Academic Press, Trondheim, Norway, 2009, 586 pp. *Polar Biology*, *34*(8), 1253–1254. https://doi.org/10.1007/s00300-011-0985-2

Bru, D., Martin-Laurent, F., & Philippot, L. (2008). Quantification of the Detrimental Effect of a Single Primer-Template Mismatch by Real-Time PCR Using the 16S rRNA Gene as an Example. *Applied and Environmental Microbiology*, *74*(5), 1660–1663. https://doi.org/10.1128/AEM.02403-07

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J., & Wittwer, C. T. (2009). The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clinical Chemistry*, *55*(4), 611–622. https://doi.org/10.1373/clinchem.2008.112797

Cohen, J. H., Berge, J., Moline, M. A., Sørensen, A. J., Last, K., Falk-Petersen, S., Renaud, P. E., Leu, E. S., Grenvald, J., Cottier, F., Cronin, H., Menze, S., Norgren, P., Varpe, Ø., Daase, M., Darnis, G., & Johnsen, G. (2015). Is Ambient Light during the High Arctic Polar Night Sufficient to Act as a Visual Cue for Zooplankton? *PLOS ONE*, *10*(6), e0126247. https://doi.org/10.1371/journal.pone.0126247

Cohen, J. H., Last, K. S., Charpentier, C. L., Cottier, F., Daase, M., Hobbs, L., Johnsen, G., & Berge, J. (2021). Photophysiological cycles in Arctic krill are entrained by weak midday twilight during the Polar Night. *PLOS Biology*, *19*(10), e3001413. https://doi.org/10.1371/journal.pbio.3001413

Connan-McGinty, S., Banas, N. S., Berge, J., Cottier, F., Grant, S., Johnsen, G., Kopec, T. P., Porter, M., & McKee, D. (2022). Midnight Sun to Polar Night: A Model of Seasonal Light in the Barents Sea. *Journal of Advances in Modeling Earth Systems*, *14*(10), e2022MS003198. https://doi.org/10.1029/2022MS003198

Dąbrowska, A. M., Wiktor, J. M., Merchel, M., & Wiktor, J. M. (2020). Planktonic Protists of the Eastern Nordic Seas and the Fram Strait: Spatial Changes Related to Hydrography During Early Summer. *Frontiers in Marine Science*, *7*, 557. https://doi.org/10.3389/fmars.2020.00557

Fuhrman, J. A., Schwalbach, M. S., & Stingl, U. (2008). Proteorhodopsins: An array of physiological roles? *Nature Reviews Microbiology*, *6*(6), 488–494. https://doi.org/10.1038/nrmicro1893

Godhe, A., Asplund, M. E., Härnström, K., Saravanan, V., Tyagi, A., & Karunasagar, I. (2008). Quantification of Diatom and Dinoflagellate Biomasses in Coastal Marine Seawater Samples by Real-Time PCR. *Applied and Environmental Microbiology*, *74*(23), 7174–7182. https://doi.org/10.1128/AEM.01298-08

Gómez-Consarnau, L., Akram, N., Lindell, K., Pedersen, A., Neutze, R., Milton, D. L., González, J. M., & Pinhassi, J. (2010). Proteorhodopsin Phototrophy Promotes Survival of Marine Bacteria during Starvation. *PLoS Biology*, *8*(4), e1000358. https://doi.org/10.1371/journal.pbio.1000358

Gorai, T., Katayama, T., Obata, M., Murata, A., & Taguchi, S. (2014). Low blue light enhances growth rate, light absorption, and photosynthetic characteristics of four marine phytoplankton species. *Journal of Experimental Marine Biology and Ecology*, *459*, 87–95. https://doi.org/10.1016/j.jembe.2014.05.013

Gouy, M., Guindon, S., & Gascuel, O. (2010). SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Molecular Biology and Evolution*, *27*(2), 221–224. https://doi.org/10.1093/molbev/msp259

Govorunova, E. G., Sineshchekov, O. A., Li, H., & Spudich, J. L. (2017). Microbial Rhodopsins: Diversity, Mechanisms, and Optogenetic Applications. *Annual Review of Biochemistry*, *86*(1), 845–872. https://doi.org/10.1146/annurev-biochem-101910-144233

Graff, J., Westberry, T., Milligan, A., Brown, M., Dall'Olmo, G., Reifel, K., & Behrenfeld, M. (2016). Photoacclimation of natural phytoplankton communities. *Marine Ecology Progress Series*, *542*, 51–62. https://doi.org/10.3354/meps11539

Guo, Z., Zhang, H., & Lin, S. (2014). Light-Promoted Rhodopsin Expression and Starvation Survival in the Marine Dinoflagellate Oxyrrhis marina. *PLoS ONE*, *9*(12), e114941. https://doi.org/10.1371/journal.pone.0114941

Han, B.-P. (2000). Effect of photoinhibition on algal photosynthesis: A dynamic model. *Journal of Plankton Research*, *22*(5), 865–885. https://doi.org/10.1093/plankt/22.5.865

Hop, H., Kovaltchouk, N. A., & Wiencke, C. (2016). Distribution of macroalgae in Kongsfjorden, Svalbard. *Polar Biology*, *39*(11), 2037–2051. https://doi.org/10.1007/s00300-016-2048-1

Huang, H.-L., Shao, Q.-W., Zhu, X.-J., Luo, J., Meng, R., Zhou, C.-X., Zhu, P., Zhu, Y.-F., & Yan, X.-J. (2019). Distribution of Karlodinium veneficum in the coastal region of Xiangshan Bay in the East China Sea, as detected by a real-time quantitative PCR assay of ribosomal ITS sequence. *Harmful Algae*, *81*, 65–76. https://doi.org/10.1016/j.hal.2018.12.001

Ishii, K., & Fukui, M. (2001). Optimization of Annealing Temperature To Reduce Bias Caused by a Primer Mismatch in Multitemplate PCR. *Applied and Environmental Microbiology*, *67*(8), 3753–3755. https://doi.org/10.1128/AEM.67.8.3753-3755.2001

Jalkanen, A. L., Coleman, S. J., & Wilusz, J. (2014). Determinants and implications of mRNA poly(A) tail size – Does this protein make my tail look big? *Seminars in Cell & Developmental Biology*, *34*, 24–32. https://doi.org/10.1016/j.semcdb.2014.05.018

Karasuyama, M., Inoue, K., Nakamura, R., Kandori, H., & Takeuchi, I. (2018). Understanding Colour Tuning Rules and Predicting Absorption Wavelengths of Microbial Rhodopsins by Data-Driven Machine-Learning Approach. *Scientific Reports*, *8*(1), 15580. https://doi.org/10.1038/s41598-018-33984-w

Karrer, E. E., Lincoln, J. E., Hogenhout, S., Bennett, A. B., Bostock, R. M., Martineau, B., Lucas, W. J., Gilchrist, D. G., & Alexander, D. (1995). In situ isolation of mRNA from individual plant cells: Creation of cell-specific cDNA libraries. *Proceedings of the National Academy of Sciences*, *92*(9), 3814–3818. https://doi.org/10.1073/pnas.92.9.3814

Keck, A. (1999). Phytoplankton assemblages related to physical gradients in an arctic, glacier-fed fjord in summer. *ICES Journal of Marine Science*, *56*, 203–214. https://doi.org/10.1006/jmsc.1999.0631

Kubiszyn, A. M., Wiktor, J. M., Wiktor, J. M., Griffiths, C., Kristiansen, S., & Gabrielsen, T. M. (2017). The annual planktonic protist community structure in an ice-free high Arctic fjord (Adventfjorden, West Spitsbergen). *Journal of Marine Systems*, *169*, 61–72. https://doi.org/10.1016/j.jmarsys.2017.01.013

Kvernvik, A. C., Hoppe, C. J. M., Lawrenz, E., Prášil, O., Greenacre, M., Wiktor, J. M., & Leu, E. (2018). Fast reactivation of photosynthesis in arctic phytoplankton during the polar night [1]. *Journal of Phycology*, *54*(4), 461–470. https://doi.org/10.1111/jpy.12750

Kvernvik, A. C., Rokitta, S. D., Leu, E., Harms, L., Gabrielsen, T. M., Rost, B., & Hoppe, C. J. M. (2020). Higher sensitivity towards light stress and ocean acidification in an Arctic sea-ice-associated diatom compared to a pelagic diatom. *New Phytologist*, *226*(6), 1708–1724. https://doi.org/10.1111/nph.16501

Lami, R., Cottrell, M. T., Campbell, B. J., & Kirchman, D. L. (2009). Light-dependent growth and proteorhodopsin expression by *Flavobacteria* and SAR11 in experiments with Delaware coastal waters. *Environmental Microbiology*, *11*(12), 3201–3209. https://doi.org/10.1111/j.1462-2920.2009.02028.x

Last, K. S., Hobbs, L., Berge, J., Brierley, A. S., & Cottier, F. (2016). Moonlight Drives Ocean-Scale Mass Vertical Migration of Zooplankton during the Arctic Winter. *Current Biology*, *26*(2), 244–251. https://doi.org/10.1016/j.cub.2015.11.038

Lavaud, J., & Goss, R. (2014). The Peculiar Features of Non-Photochemical Fluorescence Quenching in Diatoms and Brown Algae. In B. Demmig-Adams, G. Garab, W. Adams Iii, & Govindjee (Eds.), *Non-Photochemical Quenching and Energy Dissipation in Plants, Algae and Cyanobacteria* (Vol. 40, pp. 421–443). Springer Netherlands. https://doi.org/10.1007/978-94-017-9032-1_20

Lekanne Deprez, R. H., Fijnvandraat, A. C., Ruijter, J. M., & Moorman, A. F. M. (2002). Sensitivity and accuracy of quantitative real-time polymerase chain reaction using SYBR green I depends on cDNA synthesis conditions. *Analytical Biochemistry*, *307*(1), 63–69. https://doi.org/10.1016/S0003-2697(02)00021-0

Leu, E., Falk-Petersen, S., & Hessen, D. O. (2007). Ultraviolet radiation negatively affects growth but not food quality of arctic diatoms. *Limnology and Oceanography*, *52*(2), 787–797. https://doi.org/10.4319/lo.2007.52.2.0787

Levialdi Ghiron, J. H., Amato, A., Montresor, M., & Kooistra, W. H. C. F. (2008). Plastid Inheritance in the Planktonic Raphid Pennate Diatom Pseudo-nitzschia delicatissima (Bacillariophyceae). *Protist*, *159*(1), 91–98. https://doi.org/10.1016/j.protis.2007.06.002

Lewis, K. M., Arntsen, A. E., Coupel, P., Joy-Warren, H., Lowry, K. E., Matsuoka, A., Mills, M. M., Van Dijken, G. L., Selz, V., & Arrigo, K. R. (2019). Photoacclimation of Arctic Ocean phytoplankton to shifting light and nutrient limitation. *Limnology and Oceanography*, *64*(1), 284–301. https://doi.org/10.1002/lno.11039

Lin, S., Zhang, H., Zhuang, Y., Tran, B., & Gill, J. (2010). Spliced leader–based metatranscriptomic analyses lead to recognition of hidden genomic features in dinoflagellates. *Proceedings of the National Academy of Sciences*, *107*(46), 20033–20038. https://doi.org/10.1073/pnas.1007246107

Ma, M., Li, H., Wang, C., Li, T., Wang, J., Yuan, H., Yu, L., Wang, J., Li, L., & Lin, S. (2023). A comparative study reveals the relative importance of prokaryotic and eukaryotic proton pump rhodopsins in a subtropical marginal sea. *ISME Communications*, *3*(1), 79. https://doi.org/10.1038/s43705-023-00292-y

Marchetti, A., Catlett, D., Hopkinson, B. M., Ellis, K., & Cassar, N. (2015). Marine diatom proteorhodopsins and their potential role in coping with low iron availability. *The ISME Journal*, *9*(12), 2745–2748. https://doi.org/10.1038/ismej.2015.74

Markager, S., Vincent, W. F., & Tang, E. P. Y. (1999). Carbon fixation by phytoplankton in high Arctic lakes: Implications of low temperature for photosynthesis. *Limnology and Oceanography*, *44*(3), 597–607. https://doi.org/10.4319/lo.1999.44.3.0597

Marquardt, M., Vader, A., Stübner, E. I., Reigstad, M., & Gabrielsen, T. M. (2016). Strong Seasonality of Marine Microbial Eukaryotes in a High-Arctic Fjord (Isfjorden, in West Spitsbergen, Norway). *Applied and Environmental Microbiology*, *82*(6), 1868–1880. https://doi.org/10.1128/AEM.03208-15

Martinez, A., Bradley, A. S., Waldbauer, J. R., Summons, R. E., & DeLong, E. F. (2007). Proteorhodopsin photosystem gene expression enables photophosphorylation in a heterologous host. *Proceedings of the National Academy of Sciences*, *104*(13), 5590–5595. https://doi.org/10.1073/pnas.0611470104

Meng, R., Zhou, C., Zhu, X., Huang, H., Xu, J., Luo, Q., & Yan, X. (2019). Critical light-related gene expression varies in two different strains of the dinoflagellate Karlodinium veneficum in response to the light spectrum and light intensity. *Journal of Photochemistry and Photobiology B: Biology*, *194*, 76–83. https://doi.org/10.1016/j.jphotobiol.2019.03.009

Moore, C. M., Suggett, D. J., Hickman, A. E., Kim, Y.-N., Tweddle, J. F., Sharples, J., Geider, R. J., & Holligan, P. M. (2006). Phytoplankton photoacclimation and photoadaptation in response to environmental gradients in a shelf sea. *Limnology and Oceanography*, *51*(2), 936–949. https://doi.org/10.4319/lo.2006.51.2.0936

Morgan-Kiss, R. M., Lizotte, M. P., Kong, W., & Priscu, J. C. (2016). Photoadaptation to the polar night by phytoplankton in a permanently ice-covered Antarctic lake: Photoadaptation during polar night transition. *Limnology and Oceanography*, *61*(1), 3–13. https://doi.org/10.1002/lno.10107

Nilsen, F., Cottier, F., Skogseth, R., & Mattsson, S. (2008). Fjord–shelf exchanges controlled by ice and brine production: The interannual variation of Atlantic Water in Isfjorden, Svalbard. *Continental Shelf Research*, *28*(14), 1838–1853. https://doi.org/10.1016/j.csr.2008.04.015

Palmer, G., & Johnsen, S. (2015). Downwelling spectral irradiance during evening twilight as a function of the lunar phase. *Applied Optics*, *54*(4), B85. https://doi.org/10.1364/AO.54.000B85

Perry, M. J., Talbot, M. C., & Alberte, R. S. (1981). Photoadaption in marine phytoplankton: Response of the photosynthetic unit. *Marine Biology*, *62*(2–3), 91–101. https://doi.org/10.1007/BF00388170

Piredda, R., Tomasino, M. P., D'Erchia, A. M., Manzari, C., Pesole, G., Montresor, M., Kooistra, W. H. C. F., Sarno, D., & Zingone, A. (2017). Diversity and temporal patterns of planktonic protist assemblages at a Mediterranean Long Term Ecological Research site. *FEMS Microbiology Ecology*, *93*(1), fiw200. https://doi.org/10.1093/femsec/fiw200

Popova, E. E., Yool, A., Coward, A. C., Aksenov, Y. K., Alderson, S. G., De Cuevas, B. A., & Anderson, T. R. (2010). Control of primary production in the Arctic by nutrients and light: Insights from a high resolution ocean general circulation model. *Biogeosciences*, *7*(11), 3569–3591. https://doi.org/10.5194/bg-7-3569-2010

Ruiz-González, M. X., & Marín, I. (2004). New Insights into the Evolutionary History of Type 1 Rhodopsins. *Journal of Molecular Evolution*, *58*(3), 348–358. https://doi.org/10.1007/s00239-003-2557-8

Russell, T. S., Coleman, M., Rath, P., Nilsson, A., & Rothschild, K. J. (1997). Threonine-89 Participates in the Active Site of Bacteriorhodopsin: Evidence for a Role in Color Regulation and Schiff Base Proton Transfer. *Biochemistry*, *36*(24), 7490–7497. https://doi.org/10.1021/bi9702871

Sakshaug, E. (2004). Primary and Secondary Production in the Arctic Seas. In R. Stein & R. W. MacDonald (Eds.), *The Organic Carbon Cycle in the Arctic Ocean* (pp. 57–81). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-18912-8_3

Schmidt, P. J., Acosta, N., Chik, A. H. S., D'Aoust, P. M., Delatolla, R., Dhiyebi, H. A., Glier, M. B., Hubert, C. R. J., Kopetzky, J., Mangat, C. S., Pang, X.-L., Peterson, S. W., Prystajecky, N., Qiu, Y., Servos, M. R., & Emelko, M. B. (2023). Realizing the value in "non-standard" parts of the qPCR standard curve by integrating fundamentals of quantitative microbiology. *Frontiers in Microbiology*, *14*, 1048661. https://doi.org/10.3389/fmicb.2023.1048661

Seuthe, L., Rokkan Iversen, K., & Narcy, F. (2011). Microbial processes in a high-latitude fjord (Kongsfjorden, Svalbard): II. Ciliates and dinoflagellates. *Polar Biology*, *34*(5), 751–766. https://doi.org/10.1007/s00300-010-0930-9

Shi, X., Li, L., Guo, C., Lin, X., Li, M., & Lin, S. (2015). Rhodopsin gene expression regulated by the light dark cycle, light spectrum and light intensity in the dinoflagellate Prorocentrum. *Frontiers in Microbiology*, *6*. https://doi.org/10.3389/fmicb.2015.00555

Slamovits, C. H., Okamoto, N., Burri, L., James, E. R., & Keeling, P. J. (2011). A bacterial proteorhodopsin proton pump in marine eukaryotes. *Nature Communications*, *2*(1), 183. https://doi.org/10.1038/ncomms1188

Smith, C. J., & Osborn, A. M. (2009). Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology: Application of Q-PCR in microbial ecology. *FEMS Microbiology Ecology*, *67*(1), 6–20. https://doi.org/10.1111/j.1574-6941.2008.00629.x

Steindler, L., Schwalbach, M. S., Smith, D. P., Chan, F., & Giovannoni, S. J. (2011). Energy Starved Candidatus Pelagibacter Ubique Substitutes Light-Mediated ATP Production for Endogenous Carbon Respiration. *PLoS ONE*, *6*(5), e19725. https://doi.org/10.1371/journal.pone.0019725

Stephens, T. G., González-Pech, R. A., Cheng, Y., Mohamed, A. R., Burt, D. W., Bhattacharya, D., Ragan, M. A., & Chan, C. X. (2020). Genomes of the dinoflagellate Polarella glacialis encode tandemly repeated single-exon genes with adaptive functions. *BMC Biology*, *18*(1), 56. https://doi.org/10.1186/s12915-020-00782-8

Svec, D., Tichopad, A., Novosadova, V., Pfaffl, M. W., & Kubista, M. (2015). How good is a PCR efficiency estimate: Recommendations for precise and robust qPCR efficiency assessments. *Biomolecular Detection and Quantification*, *3*, 9–16. https://doi.org/10.1016/j.bdq.2015.01.005

Terhaar, J., Lauerwald, R., Regnier, P., Gruber, N., & Bopp, L. (2021). Around one third of current Arctic Ocean primary production sustained by rivers and coastal erosion. *Nature Communications*, *12*(1), 169. https://doi.org/10.1038/s41467-020-20470-z

Tilzer, M. M., Elbr□chter, M., Gieskes, W. W., & Beese, B. (1986). Light-temperature interactions in the control of photosynthesis in Antarctic phytoplankton. *Polar Biology*, *5*(2), 105–111. https://doi.org/10.1007/BF00443382

Torres-Valdés, S., Tsubouchi, T., Bacon, S., Naveira-Garabato, A. C., Sanders, R., McLaughlin, F. A., Petrie, B., Kattner, G., Azetsu-Scott, K., & Whitledge, T. E. (2013). Export of nutrients from the Arctic Ocean: ARCTIC OCEAN NUTRIENT EXPORTS. *Journal of Geophysical Research: Oceans*, *118*(4), 1625–1644. https://doi.org/10.1002/jgrc.20063

Vader, A., Marquardt, M., Meshram, A. R., & Gabrielsen, T. M. (2015). Key Arctic phototrophs are widespread in the polar night. *Polar Biology*, *38*(1), 13–21. https://doi.org/10.1007/s00300-014-1570-2

Vader, A., Nielsen, H., & Johansen, S. (1999). In vivo expression of the nucleolar group I intron-encoded I-DirI homing endonuclease involves the removal of a spliceosomal intron. *The EMBO Journal*, *18*(4), 1003–1013. https://doi.org/10.1093/emboj/18.4.1003

Valiente Moro, C., Crouzet, O., Rasconi, S., Thouvenot, A., Coffe, G., Batisson, I., & Bohatier, J. (2009). New Design Strategy for Development of Specific Primer Sets for PCR-Based Detection of *Chlorophyceae* and *Bacillariophyceae* in Environmental Samples. *Applied and Environmental Microbiology*, *75*(17), 5729–5733. https://doi.org/10.1128/AEM.00509-09

Valle, K. C., Nymark, M., Aamot, I., Hancke, K., Winge, P., Andresen, K., Johnsen, G., Brembu, T., & Bones, A. M. (2014). System Responses to Equal Doses of Photosynthetically Usable Radiation of Blue, Green, and Red Light in the Marine Diatom Phaeodactylum tricornutum. *PLoS ONE*, *9*(12), e114211. https://doi.org/10.1371/journal.pone.0114211

Vonnahme, T. R., Persson, E., Dietrich, U., Hejdukova, E., Dybwad, C., Elster, J., Chierici, M., & Gradinger, R. (2021). Early spring subglacial discharge plumes fuel under-ice primary production at a Svalbard tidewater glacier. *The Cryosphere*, *15*(4), 2083–2107. https://doi.org/10.5194/tc-15-2083-2021

Wacker, A., Piepho, M., & Spijkerman, E. (2015). Photosynthetic and fatty acid acclimation of four phytoplankton species in response to light intensity and phosphorus availability. *European Journal of Phycology*, *50*(3), 288–300. https://doi.org/10.1080/09670262.2015.1050068

Wietz, M., Bienhold, C., Metfies, K., Torres-Valdés, S., Von Appen, W.-J., Salter, I., & Boetius, A. (2021). The polar night shift: Seasonal dynamics and drivers of Arctic Ocean microbiomes revealed by autonomous sampling. *ISME Communications*, *1*(1), 76. https://doi.org/10.1038/s43705-021-00074-4

Wutkowska, M., Vader, A., Logares, R., Pelletier, E., & Gabrielsen, T. M. (2023). Linking extreme seasonality and gene expression in arctic marine protists. *bioRxiv*, 2021.11.11.467955. https://doi.org/10.1101/2021.11.11.467955

Yoshizawa, S., Azuma, T., Kojima, K., Inomura, K., Hasegawa, M., Nishimura, Y., Kikuchi, M., Armin, G., Miyashita, H., Ifuku, K., Yamano, T., Marchetti, A., Fukuzawa, H., Sudo, Y., & Kamikawa, R. (2022). *Proton-pumping rhodopsins in marine diatoms* [Preprint]. Microbiology. https://doi.org/10.1101/2022.01.18.476826

Zhang, H., Campbell, D. A., Sturm, N. R., & Lin, S. (2009). Dinoflagellate Spliced Leader RNA Genes Display a Variety of Sequences and Genomic Arrangements. *Molecular Biology and Evolution*, *26*(8), 1757–1771. https://doi.org/10.1093/molbev/msp083

Zhang, H., Nulick, K. J., Burris, Z., Pierce, M., Ma, M., & Lin, S. (2022). *Dinoflagellate Proton-Pump Rhodopsin Gene in Long Island Sound: Diversity and Spatiotemporal Distribution* [Preprint]. Molecular Biology. https://doi.org/10.1101/2022.08.24.505117

Zhang, Y., Lin, X., Shi, X., Lin, L., Luo, H., Li, L., & Lin, S. (2019). Metatranscriptomic Signatures Associated With Phytoplankton Regime Shift From Diatom Dominance to a Dinoflagellate Bloom. *Frontiers in Microbiology*, *10*, 590. https://doi.org/10.3389/fmicb.2019.00590

Supplementary material

Protocols

1. Extractions and purification.

1.1. DNA Extraction

Total DNA was extracted with the DNeasy® Plant Pro Kit (50) from QIAGEN.

Materials

Reagents:

- 96% Ethanol.
- Reagents provided in kit.
- Molecular biology grade zirconium beads.

Equipment:

- Heating block.
- Electronic rotor-stator homogenizers.
- Microcentrifuge.

Sample Prerequisites

- All samples were stored at -80°C in 2 mL cryo tubes.
- Samples were thawed on ice.

Notes before starting:

- All centrifugation steps must be performed at room temperature.
- Buffers AW2 and Buffer AW1 are supplied as concentrates. Before using for the first time, add the appropriate amount of ethanol (96-100%) as indicated on the bottle to obtain a working solution.
- If Buffer AW1 forms precipitates, warm to 65°C to redissolve.

Protocol:

1. Divide zirconium beads (200 µm, molecular grade, OPS diagnostics) from one bead tube into two 1.5mL Eppendorf ® tubes (1 needed per sample).
2. Cut filter halves into 4-5 small pieces using disposable scalpel and Petri dish. Transfer filter pieces to an Eppendorf ® tube with beads and add 400 µl Buffer AP1.
3. Beat the samples for 1' at 30 Hz in a bead-beater (store blocks in the fridge).
4. Reverse the position of the tubes within the adaptor set (to prevent between sample variation in homogenization) and repeat bead beating (1' at 30 Hz).
5. Quick spin tubes to remove foam (~30'' using microcentrifuge).
6. Pipette the mixture, avoiding filter and beads, into a new 1.5 mL Eppendorf ® tube.
7. Re-extract the filter: Add 400 µl Buffer AP1 to the filter and repeat steps 3-6, adding the mixture to the same tube as step 6 (~800 µl total).
8. Add 4 µl RNaseA to the tubes with filter extractions.
9. Incubate all tubes for 10' at 65°C. **Vortex** 2-3 times during incubation (for cell lysis).
10. Add 260 µl Buffer P3 to the lysate, invert to mix, and incubate **on ice** for 5' (precipitation of detergent, proteins and polysaccharides).
11. Centrifuge the lysate for 5' at 20,000g.
12. Pipet 675 µl of **supernatant** into a QIAshredder Mini spin column placed in a 2 mL collection tube.
13. Centrifuge for 2' at 20,000g. ***NB: DNA is now in the flow-through.***
14. Transfer the **flow-through** into two new 1.5 mL Eppendorf ® tubes. Do not disturb the pellet which may occur in the collection tube. *Typically, 2 x 400-450 µl of lysate is recovered per sample.*
15. Repeat steps 12 and 13 with the remaining supernatant.
16. Add 1.5 times the lysate volume of Buffer AW1 to the cleared lysate, and mix by inverting the tube a few times.
17. Pipet 650 µl of the mixture from step 16, including any precipitate that may have formed, into a DNeasy Mini spin column placed in a 2 mL collection tube. Use one spin column for the 2 tubes from the same sample. Centrifuge for 1' at ≥6000*g* and discard flow-through. ***NB: DNA is now attached to the filter of the spin column.***

18. Repeat step 17 with remaining sample (from both tubes) until all has been used on one spin column. Discard final flow-through. ***All the DNA is now attached to the filter of the spin column.***
19. Add 500 μl Buffer AW2 to the DNeasy Mini spin column. Centrifuge for 1' at ≥6000*g*. Discard flow through. ***NB:*** Here we are rinsing the filter, ***the DNA stays on the filter.***
20. Add another 500 μl Buffer AW2 to the DNeasy Mini Spin column and centrifuge for 1' at ≥6000*g*. Discard the flow-through. NB: continuation of rinsing the filter.
21. Centrifuge the DNeasy Mini spin column for 2' at 20,000g to dry the membrane (to remove residual ethanol from the column/sample).
22. Transfer the DNeasy Mini spin column to a 1.5 mL Eppendorf ® tube (prelabelled with extraction date, sample depth, location, date and fraction for long-term storage).
23. Pipette 75 μl Buffer AE directly onto the DNeasy membrane. **Incubate for 5' at RT**. Centrifuge for 1' at ≥6000*g* to elute. ***NB: DNA is now in the flow-through.***
24. Repeat step 22 for a total of 150 μl of elute with DNA for each sample.

1.2 RNA Extraction

Total RNA was extracted with the RNAqueous Total RNA Isolation Kit (Invitrogen, Thermo Scientific ®):

Materials

Reagents:

- 64% Ethanol.
- Reagents provided in kit.
- Molecular biology grade zirconium beads.

Equipment:

- Heating block.
- Ice bath.
- Electronic rotor-stator homogenizers.
- Microcentrifuge.

Sample Prerequisites

- All samples were stored at -80°C in 2 mL cryo tubes with 600µl Lysis/Binding Solution (Thermo Scientific®).
- Samples must be thawed on ice, essential for preventing ice crystals from forming and rupturing cellular compartments.

Notes before starting:

- Label two 2 mL and one 1.5 mL Eppendorf ® tube per RNA sample.
- Divide beads that are in one bead-tube evenly between two 2 mL Eppendorf ® tubes.
- Add 64 mL of 100% EtOH to the Wash Solution #2/3 Concentrate. Mix thoroughly.
- Prewarm an aliquot of Elution Buffer (60 µl per sample) on 75° heating block.

Protocol:

1. Thaw samples on ice, then vortex.
2. Transfer **supernatant** to 2 mL Eppendorf ® tubes with beads.
3. Add 600 µl LB buffer to the filter cartridge, vortex and transfer **supernatant** to a new 2 mL Eppendorf ® tube with beads.
4. "Beat beating" at 2 x 1/22 s.
5. Add 600 µl 64% EtOH to each tube. Mix gently by turning the tube upside down a few times.
6. Place filter cartridge (from kit) into a labelled collection tube (from kit) and add 600 µl Lysate-EtOH.
7. Spin for 1' at 13.000g. Discard **flow-through.**
8. Add 600 µl Lysate-EtOH and repeat spin until all liquid from both 2 mL tubes has been passed through one membrane (one membrane per sample).
9. Discard **flow-through**.
10. Wash with 700 µl Wash 1. Spin for 1' at 13.000g.
11. Discard **flow-through.**
12. Wash with 500 µl Wash 2/3.
13. Repeat step #12.

14. Discard **flow-through**. Spin for 1' to ensure that the membrane is completely dry. Inspect cartridge to ensure there are no residual traces of wash solution (this may inhibit downstream steps).
15. Transfer filter cartridge to labelled 1.5 mL Eppendorf ® tube. Elute by adding 50 µl of prewarmed Elution Buffer.
16. Spin for 30 seconds at 13.000g.
17. Repeat elution with 10 µl of prewarmed Elution Buffer.
18. Store RNA at -80°C.

1.3 DNase treatment

Materials

Reagents:

- Reagents provided in kit.

Equipment:

- Centrifuge/ microcentrifuge.
- Vortex.
- Qubit.

Protocol:

1. Take 5 µl of combined RNA for agarose gel electrophoresis.
2. Take 54 µl of RNA for DNAse treatment.
3. Mix 54 µl RNA with:
   a) 6 µl 10x DNAse buffer
   b) 2 µl TURBO DNAse buffer
4. Mix gently by pipetting and incubate at room temperature (RT) for 20'.
5. Vortex DNAse inactivation reagent and add 6 µl, mix gently by pipetting.
6. Incubate at RT for 5', mix occasionally by flicking the tube.
7. Spin at 10,000 rcf for 1.5'.
8. Set the pipette to 62 µl and transfer the supernatant (without disturbing the pellet) into labelled 1.5 mL Eppendorf ® tubes.

Afterwards:

- Use 5 µl of DNAse-treated RNA for agarose gel electrophoresis (method below).
- Take 2 µl of DNAse-treated RNA and check on Qubit.
- Take 5 µl of DNAse-treated RNA for cDNA synthesis and PCR.
- Freeze remaining product at -80°C.

Agarose gel electrophoresis:

1. Prepare a 0.7% agarose gel using RNA electrophoresis chamber (Agarose: 0.7g, TA buffer: 100 mL, and GelRed®: 3 µl).
2. Prepare PCR strip with 1 µl loading dye per sample. Add 5 µl of RNA or 5 µl of DNAse-treated RNA.
3. Run 3 µl FastRuler Low Range ladder for size indentification.

Invitrogen™ Qubit™ Fluorometer:

1) Prepare standards and RNA for Qubit (enough working solution for 200 µl per sample/ standard). NB: light sensitive.
   a) Standard 1: 190 µl working solution +10 µl ST1
   b) Standard 2: 190 µl working solution + 10 µl ST2
   c) Samples: 198 µl working solution + 2 µl RNA

1.4 Denaturation of RNA

Materials:

- 1µl RNA (DNA-se treated)
- 1µl Random Hexamer primer (100µM, Thermo Scientific®) or poly-A primer (50 µM)
- 1µl dNTP mix (10mM each)
- 10µl H_2O, nuclease-free (Ambion)

Protocol:

- Mix with and spin-down.

- Incubate at 65°C for 5’ and put directly on ice.

1.5 cDNA synthesis

cDNA synthesis was performed using the Thermo Scientific™ SuperScript™ IV Reverse Transcriptase on denatured RNA.

Materials:

- 4 µl 5x Superscript IV buffer
- 1µl 0.1M DTT
- 1µl RNAse inhibitor (40U/µl, Thermo Scientific™)
- 1µl Superscript IV RT (200U/µl, Invitrogen)
- 13µl denatured DNAse-treated RNA

The protocol used for cDNA synthesis is determined by which primer is being used. Once all reagents and genetic material have been combined and spun down, Random hexamer primers require the following protocol: 23°C for 10’, 50°C for 10’ and 80°C for 10’ whereas Poly-A primers require only two steps: 50°C for 10’ and 80°C for 10’. Store material at -20°C.

2. Reaction Mixtures

2.1 PCR

Notes:

- Once the Mastermixes were made in 1.5 mL Eppendorf ® tubes, they were spun down and kept in an Eppendorf Tube Cooling Stand until use.
- MicroAmp™ 8-Tube Strips of 0.2 mL with coned caps were used.
- PCR strips were prepared with cDNA and Mastermix on a PCR-Cooler Tube Rack which was otherwise stored at -5°C.
- PCR strips were spun down prior to reaction.

For 1 replicate:

Material	**Quantity (µl)**
DreamTaq Buffer	2.5
dNTPs	2
Fp	0.5
Rp	0.5
DreamTaq	0.25
Milli-Q H_2O	18.5
DNA/cDNA	1

2.2 qPCR

Notes:

- SYBR® Green Mastermix (2x) must be heated at 95°C for 10'to activate the enzyme.
- Once Mastermix with primers was made, they were kept in an Eppendorf Tube Cooling Stand with a lid on top to prevent damage from ambient light whilst qPCR plate was loaded.
- MicroAmp 96 well 0.1mL plates were loaded with cDNA and qPCR mastermix on a PCR-Cooler Tube Rack which was otherwise stored at -5°C.
- Plates were sealed with Optical Sealing films for qPCR.

For 1 replicate:

Material	**Quantity (µl)**
SYBR® Green Mastermix (2x)	10
Nuclease-free water	8.6
Fp	0.2
Rp	0.2
cDNA	1

3. Product purification

SPRI beads

Materials

Reagents:

- Freshly made SPRI bead solution (instructions below).
- 80% Ethanol (EtOH), non-denatured (prepared fresh for optimal results).
- Milli-Q (or standard buffer solution such as TE) for DNA elution.
- GeneRuler Low-Range Ladder (Thermo Scientific™) or similar variant.

Equipment:

- Magnetic separation rack.
- MiniSpin ® Centrifuge.

Sample Prerequisites

- Samples should be newly-made PCR products.
- Sample volume should be ≥ 40 µl, a lower volume will decrease pipetting accuracy of SPRI beads.
- DNA fragments within PCR product should be size selected in a range no smaller than 150bp.

Notes before starting:

- SPRI bead solution should be vortexed for ~15-20 seconds before use so that they are suspended in solution and not sticking to the bottom of the tube.
- SPRI bead solution should be at room temperature before use.
- Prepare standard tube rack on working bench (samples go back and forth between magnetic stand and tube rack).
- Depending on the ratio being used, samples should be in 1.5-2 ml Eppendorf ® tubes.
- Sterile Eppendorf ® tubes should be ready for placing the cleaned samples into.

SPRI bead solution:

In a 50 mL Falcon Tube: add and mix the following ingredients:

1. 1% beads (SeraMag Speed Beads- Thermo Scientific™).
2. 18% PEG-8000.
3. 1M NaCl.
4. 10mM Tris (pH 8.0).
5. 1mM EDTA (pH 8.0).

Protocol:

1. Add beads to product and mix by pipetting a few times. Flick the Eppendorf ® tube for ~15-20 seconds, then MiniSpin ® Centrifuge for ~10-15 seconds.
2. Incubate at room temperature (RT) for 5' (5 minutes) on tube rack.
3. After 5 minutes, MiniSpin ® Centrifuge briefly for ~10 seconds.
4. Place tubes on magnetic stand for 2'.
5. With the tubes **still on the magnetic stand**, remove supernatant (liquid above) and close lids on tubes as you go, if not the samples will dry out.
6. Add 100 µl 80% EtOH and mix by pipetting (NB still on magnetic rack).
7. Incubate for 1'.
8. Remove supernatant and leave tubes open at RT for 2' to allow EtOH evaporate (NB keep an eye on the products in the tubes, do not let them dry out completely).
9. Elute by **removing the tubes** from magnetic stand and add 10 µl MilliQ-H_2O. Mix by flicking for ~15-20 seconds.
10. Incubate for 2' at RT on tube rack with lids closed- **not on magnetic rack**.
11. Place on magnet rack for 2'.
12. Keep on magnet whilst removing the supernatant (**this is your cleaned product**) and transfer to sterile Eppendorf ® tubes.
13. If unsure that the cleaning has worked, pipet out a small amount of the product and verify on a gel.

4. Gel extraction

Gel extraction was performed using the QIAquick Gel Extraction kit from ©QIAGEN according to manufacturer's recommendations:

Notes before starting:

- Add ethanol (96–100%) to Buffer PE before use (see bottle label for volume).
- All centrifugation steps are carried out at 17,900 x g (13,000 rpm) in a conventional table-top microcentrifuge at room temperature (15–25°C).

Method:

1. Excise the DNA fragment from the agarose gel with a clean, sharp scalpel. Minimize the size of the gel slice by removing extra agarose.
2. Weigh the gel slice in a colorless tube. Add 3 volumes of Buffer QG to 1 volume of gel (100 mg, or approximately 100 µl).
3. Incubate at 50°C for 10 min or until the gel slice has completely dissolved. To help dissolve gel, mix by vortexing the tube every 2–3 min during the incubation. **IMPORTANT**: Solubilize agarose completely. For >2% gels, increase incubation time.
4. After the gel slice has dissolved completely, check that the color of the mixture is yellow (similar to Buffer QG without dissolved agarose). If the color of the mixture is orange or violet, add 10 µl of 3 M sodium acetate, pH 5.0, and mix. The color of the mixture will turn to yellow. The adsorption of DNA to the QIAquick membrane is efficient only at pH ≤7.5. Buffer QG contains a pH indicator that is yellow at pH ≤7.5 and orange or violet at higher pH, allowing easy determination of the optimal pH for DNA binding.
5. Add 1 gel volume of isopropanol to the sample and mix. For example, if the agarose gel slice is 100 mg, add 100 µl isopropanol. This step increases the yield of DNA fragments ≤500 bp and ≥4 kb. For DNA fragments between 500 bp and 4 kb, addition of isopropanol has no effect on yield. Do not centrifuge the sample at this stage.
6. Place a QIAquick spin column in a provided 2 ml collection tube.
7. To bind DNA, apply the sample to the QIAquick column, and then centrifuge for 1 min.
8. Discard flow-through and place QIAquick column back into the same collection tube. Collection tubes are reused to reduce plastic waste. **Recommended**: Add 0.5 ml of Buffer QG to QIAquick column and centrifuge for 1 min. This step will remove all traces of agarose. This is only required if the DNA will be used for direct sequencing, in vitro transcription or microinjection.

9. To wash, add 0.75 ml of Buffer PE into the QIAquick column and centrifuge for 1 min. **Note**: If the DNA will be used for salt-sensitive applications, such as blunt-end ligation and direct sequencing, let the column stand 2–5 min after addition of Buffer PE before centrifuging.

5. Bacterial cloning

Bacterial cloning was performed using both the CloneJet PCR Cloning Kit and TransformAid Bacterial Transformation Kit from Thermo Scientific™ according to manufacturer's recommendations.

1. Sticky End Cloning Protocol

- For cloning PCR products with 3'dA overhangs generated by Taq DNA polymerase or enzyme mixtures containing Taq DNA polymerase.
- Use in a 3:1 molar ratio with pJET1.2/blunt PCR product.

1. Set up the blunting reaction on ice by combining the following:

Component	Volume
2X Reaction Buffer	10 µl.
PCR product	7 µl (0.15 pmol ends)
Water (nuclease free)	0 µl (to 18 uL total volume)
DNA Blunting Enzyme	1 µl
Total volume	*18* µl

2. Vortex briefly and centrifuge for 3-5''.
3. Incubate the mixture at 70 °C for 5'. Chill on ice.
4. Set up the ligation reaction on ice by adding the following to the blunting reaction mixture:

Component	Volume
pJET1.2/blunt Cloning Vector (50 ng/µL)	1 µl (0.05 pmol ends)

T4 DNA Ligase	1 µl
Total volume	*20* µl

1. Vortex briefly and centrifuge for 3-5'' to collect drops.
2. Incubate the ligation mixture at RT for 5'.
3. Use the ligation mixture directly for transformation. Store ligation mixture at -20°C if transformation is postponed. Thaw on ice and mix carefully before transformation.

2. Cloning: transformation

(TransformAid Bacterial Transformation Kit, Thermo Scientific™). Each experiment should include a positive (i.e. pUC19 of known concentration) and negative (no DNA) control.

1. Plate JM107 (from glycerol freezer stocks) on LB plate. Grow ON at 37°C.
2. Inoculate 2 ml of LB medium with 1 colony of JM107 (from plate, plate can be kept at 4°C for max 10 days, should be transferred to new plate after a week). Grow ON at 37°C (in Nunc tube with rotation).
3. On day of transformation, preheat 1.5 ml of C-medium and LB plates (with 50 µg/ml ampicillin, one plate for each transformation + two extra for positive and negative controls) at 37°C for at least 20'. Thaw T-solutions A and B, mix contents thoroughly (B needs vortexing). Combine 250 µL A and 250 µL B, keep T-solution on ice at all times. (1.5 ml of C-medium and 500 µL T-solution is enough for two transformation reactions).
4. Add 150 µL of overnight culture to 1.5 ml of preheated C-medium. Incubate 20' at 37°C (nunc tube or tape on 2 ml epp tube, rotation).
5. Pellet cells by centrifugation on 6000 rpm, 1'. Remove supernatant.
6. Resuspend cells in 300 µL T-solution by gently pipetting up and down (keep cells on ice, difficult to get cells resuspended). Incubate on ice for 5'.
7. 6000 rpm, 1'. Remove supernatant.
8. Resuspend cells in 120 µL T-solution by **gently** pipetting up and down. Incubate 5' on ice.
9. Add ligation mix (5 µL) or control DNA (2 µL of pUC19 10 pg/µL) into marked epp tubes. Chill on ice for at least 2'.

10. Add 50 µL of competent cells to each of the DNA tubes, mix **gently** (!). Incubate on ice for 5'.
11. Plate immediately on the pre-warmed LB plates. Leave ON at 37°C. Plate left-over (untransformed) cells in T-solution onto separate plate (= negative control).
12. Calculate transformation efficiency for your competent cells using the concentration of the positive control. Should be $>10^7$ cpu/µg DNA.

3. Colony PCR

To test whether the colonies have inserts of the expected size, single colonies are picked using a clean pipette tip and used directly in PCR reactions. Label colonies prior to picking, either by re-plating onto pre-labelled Amp-containing LB plates, or by numbering colonies on underside of the plate you are picking from.

1. Set up a PCR mastermix of the following (make enough to include negative control):

Water	18.6 µl
dNTP mix (2.5 mM each)	2.0 µl
10x DreamTaq buffer	2.5 µl
pJet1.2F primer (10 mM)	0.25 µl
pJet1.2R primer (10 mM)	0.25 µl
DreamTaq (5 U/µl)	0.2 µl
Total	*25.0 µl*

3. Mix well, spin down, divide in tubes. Add colony to tube, mix, spin down to get rid of bubbles.
4. PCR cycle: 95°C-5'; 25x (95°C-30''; 60°C-30''; 72°C -1'30''); 72°C-5'; 10°C-soak
5. Check results on 0.7% agarose gel (1xTAE). Use 2 µl low range ladder as size control.

Primer sequences:

pJet1.2F 5’- CGACTCACTATAGGGAGAGCGGC -3’

pJet1.2R 5’- AAGAACATCGATTTTCCATGGCAG -3’

4 Standard Curve

Equipment

5. Freezer stocks of plasmid
6. Monarch® Plasmid Miniprep Kit from NEB®
7. Monarch® PCR & DNA Cleanup Mit from NEB®
8. Fastdigest HindIII Kit from ThermoScientific™

Protocol

1. Prepare an overnight culture (12-16h) from freezer plasmid stock in 2 mL LB media+Amp with shaking at 37°C.

2. Prepare plasmid with Monarch® Plasmid Miniprep Kit:
    a. Pellet bacterial culture by centrifugation: 30sec at 13 000RPM
    b. Resuspend pellet in 200 µl plasmid resuspension buffer (B1) by pipetting up and down.
    c. Add 200 µl Plasmid lysis buffer (B2). Invert tube 5-6 times and incubate at room temperature for 1 minute. Solution becomes transparent and dark pink.
    d. Add 400 µl of Plasmid Neutralization Buffer (B3). Invert tube until color is uniformly yellow and precipitate forms. Incubate at room temperature for 2 minutes.
    e. Centrifuge: 5 minutes at 13 000RPM. Pellet should become compact.
    f. Transfer supernatant to spin column.
    g. Centrifuge: 1 minute at 13 000RPM. Discard flow-through
    h. Add 200 µl Plasmid wash buffer 1.
    i. Centrifuge: 1 minute at 13 000RPM. Discard flow-through

j. Add 400 µl Plasmid wash buffer 2.
k. Centrifuge: 1 minute at 13 000RPM.
l. Transfer column to a clean 1.5 mL Eppendorf tube. NB: do not come in contact with flow-through!
m. Add 30 µl DNA elution buffer to the center of the matrix.
n. Wait 1 minute then spin for 1 minute at 13 000RPM.

3. Cut plasmid with Fastdigest HindIII kit
    a. Combine the following reaction components at room temperature:
        i. 145 µl MilliQ
        ii. 25 µl DNA
        iii. 20 µl 10x Fastdigest buffer
        iv. 10 µl Fastdigest enzyme
    b. Mix gently and spin down
    c. Incubate for 15 minutes at 37°C in heatblock.
    d. Incubate for 10 minutes at 80°C.

9. Clean-up with Monarch® PCR & DNA Cleanup kit
    a. Dilute sample by adding 400 µl DNA cleanup binding buffer. Mix by pipetting up and down.
    b. Insert column into collection tube and load the sample onto the column.
    c. Centrifuge: 1 minute at 13 000RPM. Discard flow-through
    d. Add 200 µl DNA wash buffer.
    e. Centrifuge: 1 minute at 13 000RPM. Discard flow-through
    f. Add 200 µl DNA wash buffer.
    g. Centrifuge: 1 minute at 13 000RPM. Discard flow-through
    h. Transfer column to a clean 1.5 mL Eppendorf tube. NB: do not come in contact with flow-through!
    i. Add 20 µl DNA elution buffer to the center of the matrix.
    j. Wait 1 minute then spin for 1 minute at 13 000RPM.

10. Check quantity and purity of your stock solution:
    a. Mix 0.5 µl of DNA with 4.5 µl MilliQ and run on a 0.7% gel.
    b. Check concentration on nanodrop (use elution buffer as a blank).

11. Make dilution series
    a) Dilute an aliquot (i.e. 3 µl) to a copy number of 10^9 for your stock solution.
    b) Make 1:10 serial dilutions (enough to run triplicates).

Physicochemical data

Table S1. Nutrient analysis data from time series for Silicate, Phosophate and Nitrate/Nitrite.

Date	Replicate	SiO_2µmol/L	NOxµmol/L	PO_4µmol/L	NO_3-Nµmol/L
17/01/2023	1	3,682	3,871	0,566	3,939
	2	3,615	3,822	0,566	3,9
01/02/2023	1	0	1,698	0,42	1,735
	2	2,043	3,999	0,544	4,036
15/02/2023	1	2,097	4,081	0,544	4,116
	2	1,731	3,892	0,565	3,925
01/03/2023	1	2,9	2,819	0,377	2,926
	2	4,405	4,298	0,577	4,391
15/03/2023	1	4,518	4,894	0,59	4,957
	2	4,052	4,512	0,538	4,572
31/03/2023	1	4,726	4,83	0,58	4,953
	2	4,796	4,849	0,567	4,972

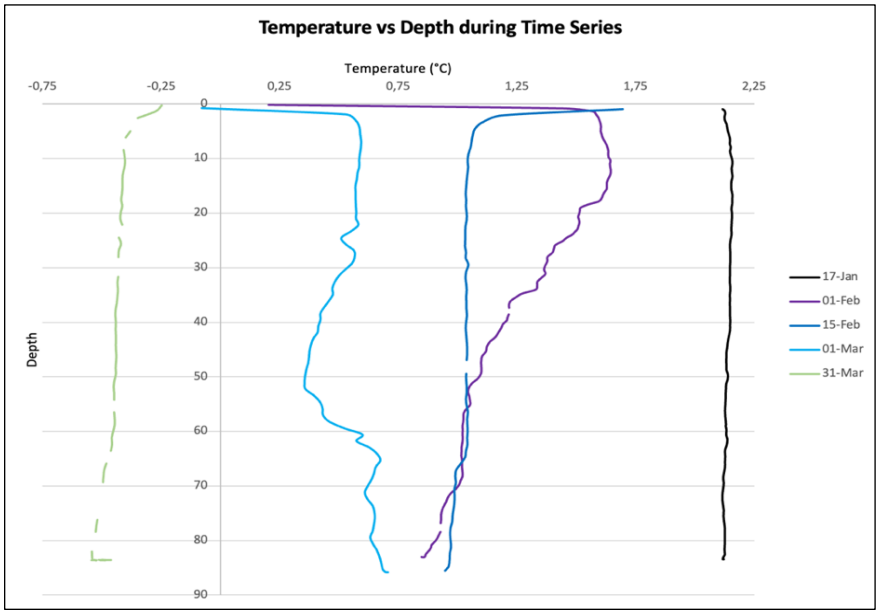


Figure S1. CTD data showing Depth plotted against Temperature throughout Time Series.

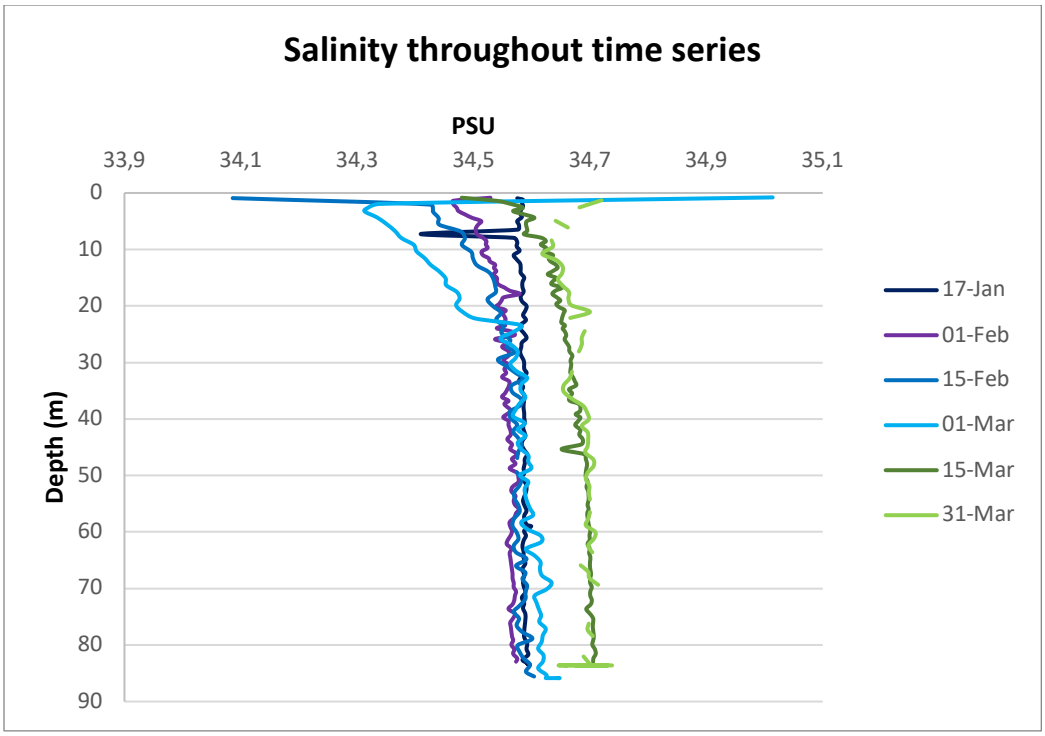


Figure S2. CTD data showing Depth plotted against Temperature throughout Time Series.

Primers

	Primer name	Sequences (5'-3')	Tm from IDT	Target organism	Target gene	Reference
	proteo F1	TBCAYTACWTSTACATGCG	50.2	*Karlodinium*	Proteorhodopsin	This study
	proteo R1	GTARAACTCVAYCATCTGCA	51.5	*Karlodinium*	Proteorhodopsin	This study
	proteo F2	TGGATGGTTCTWCATCTTGT	51.1	*Karlodinium*	Proteorhodopsin	This study
	proteo R2	ARCCRACNGTGACRATGAA	54.4	*Karlodinium*	Proteorhodopsin	This study
	blue F	ATCACHGTTCCNYTGCAGATG	56.0	*General eukaryote*	General primer for active site of Blue-light absorbing rhodopsin	This study
	uncultF1	ATCACYGTYCCYCTGCAGATGAT	59.2	*Uncultured eukaryote*	Q and S active site in blue light rhodopsin	This study
	uncult R1	CAAGAGCARGCGRWAGAACAT	55.4	*Uncultured eukaryote*	Q and S active site in blue light rhodopsin	This study
	uncultF2	CTCTACCGBTACATYGACTG	52.6	*Uncultured eukaryote*	Q and S active site in blue light rhodopsin	This study
	uncult R2	CAGSARSCRCCAGAACAT	55.5	*Uncultured eukaryote*	Q and S active site in blue light rhodopsin	This study
	uncultF3	GTCTACCGCTATGTGGACTG	54.7	*Uncultured eukaryote*	Q and S active site in blue light rhodopsin	This study
Blue absorbing rhodopsin	*uncult R3*	GAGCAAGAGCTTCCAGAACAT	55.1	*Uncultured eukaryote*	Q and S active site in blue light rhodopsin	This study
Dinoflagellate	*DinoSL-F*	CCGTAGCCATTTTGGCTCAAG	56.6 (NEB 60)	*All dinoflagellates*	Spliced leader	This study
	xantho F1	GAGATYCTBYTGGTGATGAA	51.0	*Prorocentrum*	Green light rhodopsin	This study
	xantho R1	TCGGGGAGGTTCATCACCA	58.3	*Prorocentrum*	Green light rhodopsin	This study
	xantho F2	GAYGCMTAYCGBTAYATGGA	53.2	*Prorocentrum*	Green light rhodopsin	This study
	xantho R2	TCRYYRAGCTTCATCACCA	53.8	*Prorocentrum*	Green light rhodopsin	This study
	xantho R3	CAACTCCTGCACAATGTACA	52.9	*Prorocentrum*	Green light rhodopsin	This study
	xantho R4	GAGCTCRTASACRATGTACA	51.6	*Prorocentrum*	Green light rhodopsin	This study
	green F1	TAYATGGATTGGCTCTTGAC	50.7	*Prorocentrum*	Green light rhodopsin	This study
	green F2	TACATGGACTGGTTGTTGAC	52.3	*Prorocentrum*	Green light rhodopsin	This study
	phaeo F1	AACGACGCRTACCGTTACATGGA	59.6	*Phaeocystis / Pseudo nitchzia*	Green light rhodopsin	This study
	phaeo R1	TTGTARCCAGAGATGATCAT	49.4	*Phaeocystis / Pseudo nitchzia*	Green light rhodopsin	This study
Green absorbing xanthorhodopsin	*phaeo R2*	GTYTRSACRAYGTACAGGAA	52.1	*Phaeocystis / Pseudo nitchzia*	Green light rhodopsin	This study
	rbcl- A(F)	CGTATTCGGTTTCAAAGCTATC	51.7	General eukaryote	Regulation of Rubisco activity- halotype A	Ghiron et al. 2008
	rbcl- A(R)	CATCTAAATTAGTTTGACGTAAGAT	49.2	General eukaryote	Regulation of Rubisco activity- halotype A	Ghiron et al. 2008
	rbcl-B(F)	ACGTATTCGGTTTCAAAGCTGTA	54.6	General eukaryote	Regulation of Rubisco activity- halotype B	Ghiron et al. 2008
	rbcl- B(R)	AACATCTAAATTAGTTTGACGTAAT	49.0	General eukaryote	Regulation of Rubisco activity- halotype B	Ghiron et al. 2008
	Lhcf (F)	GTTGGATTCCTTGAGCTTGC	54.3	General eukaryote	Light harvesting complex	This study
	Lhcf (R)	AGCTTGGTCTCCTCATCGAA	55.7	General eukaryote	Light harvesting complex	This study
	psbAK (F)	TTATGCACCCATTCCACATG	53.1	General eukaryote	Gene encoding the PS2 reaction centre protein in (?)	Morgan-Kiss, R.M et al. 2016
Photosynthetic Machinery	*psbAK (R)*	GATTAAACGTCCAAAGTAACCG	51.9	General eukaryote	Gene encoding the PS2 reaction centre protein in (?)	Morgan-Kiss, R.M et al. 2016
	Euk528F	CCGCGGTAATTCCAGCTC	57	General eukaryote	18S V4 Region	Godhe et al., 2008
	Diatom 18SRL	CAATGCAGWTTGATGAWCTG	50	Diatoms	18S V4 Region	Godhe et al., 2008
	Dino 18SRL	GAGCCAGATRCDCA CCCA	59	Dinoflagellates	18S V4 Region	Godhe et al., 2008
	18S V4 F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG	66	General eukaryote	18S V4 Region	Piredda et al., 2017
	18S V4 R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG	66	General eukaryote	18S V4 Region	Piredda et al., 2017
Eukaryote miscellaneaous	*1209*	CAGGTCTGTGATGCCCTT	54	General eukaryote	18S V4 Region	Giovannoni et al., 1988

Table S2. 1st set of primers designed and ordered from literature.

	Primer name	Sequences (5'-3')	Tm from IDT	Target organism	Target gene	Reference
Photosynthetic Machinery	*rbclF595*	GACTTCACCAAAGACGACGA	54.7	General eukaryote	Rubisco cycle	Morgan-Kiss, R.M et al. 2016
	rbcl R	GGAAGTGGATACCGTGGTTT	54.7	General eukaryote	Rubisco cycle	Morgan-Kiss, R.M et al. 2016
	RhbF1	ATGGCYGGATGGTTCTTCAT	55.5	General eukaryote	Blue light absorbing rhodopsin	This study
	RhbR1	GTARCCRAGWGGGTAGAT	50.4	General eukaryote	Blue light absorbing rhodopsin	This study
	RhbF2	CTGGGSHTTCATTCTGTTCGA	55.8	General eukaryote	Blue light absorbing rhodopsin	This study
	RhbR2	GTAGCCMARRGGGTAGAT	52.0	General eukaryote	Blue light absorbing rhodopsin	This study
	RhbF3	TTACATBGATTGGTCYATCAC	50.3	General eukaryote	Blue light absorbing rhodopsin	This study
Blue absorbing rhodopsin	*RhBR3*	GGCTGSACYGCAGAGAGGATC	61.1	General eukaryote	Blue light absorbing rhodopsin	This study
	DinoGF1	ACYTACCCSTTCGTGTACAT	54.4	All Dinoflagellates	Green light absorbing rhodopsins in all Dino.	This study
	Dino GR1	GARTAGCCRAYCTGCTC	51.1	All Dinoflagellates	Green light absorbing rhodopsins in all Dino.	This study
	DinoGF2	TAYCGMTAYGTSGACTGGCTG	57.5	All Dinoflagellates	Green light absorbing rhodopsins in all Dino.	This study
Dinoflagellate	*Dino GR2*	AGGTYMTCYTGGATCTCACC	54.7	All Dinoflagellates	Green light absorbing rhodopsins in all Dino.	This study

Table S3. 2nd set of primers designed and ordered from literature.

Primer name	Sequences (5'-3')	Tm from IDT	Target organism	Target gene	Reference
K-RH-F	TCCTTGAACTTGGCGTATGACTTG	57.3	*Karlodinium venificum*	Rhodopsin	Meng et al. 2019
K-RH-R	TCCTTCTCTGCGGTGGTCTTC	58.9	*Karlodinium venificum*	Rhodopsin	Meng et al. 2019
K-RBCL-F	CGGGAAAGGTTTTTATTCTCTATG	51.8	*Karlodinium venificum*	Rubisco cycle	Meng et al. 2019
K-RBCL-R	GCTTCCATTGTTGCTCCAGT	55.8	*Karlodinium venificum*	Rubisco cycle	Meng et al. 2019
K-LH-F	CAGSATACYTSGCRTTTCAGCAA	57.9	*Karlodinium venificum*	Light-harvesting center	This study
K-LH-R	TGATCRGADCGRATYGACCAAGG	58.5	*Karlodinium venificum*	Light-harvesting center	This study
K-PSB-F	CAGSATACYTSGCRTTTCAGCAA	54.3	*Karlodinium venificum*	psbA protein	This study
K-PSB-R	ATTAAGATGTTATGCTCTGCCTG	52.9	*Karlodinium venificum*	psbA protein	This study
Pr-RH-F	ACTACGGCGAACTGACTGTG	57.0	*Prorocentrum sp.*	Rhodopsin	Zhang et al. 2019
Pr-RH-R	TCGAACACAATGCACAAGAA	52.8	*Prorocentrum sp.*	Rhodopsin	Zhang et al. 2019
Pr-RH-F2	ATGWKYTTCTTCYTGTACAT	47.6	*Prorocentrum sp.*	Rhodopsin	This study
Pr-RH-R2	ACMACTGGRTAKGTGCACCA	57.6	*Prorocentrum sp.*	Rhodopsin	This study
Pr-RBCL-F	TACTACCGCCAGGAGTGGGAGGG	64.1	*Prorocentrum sp.*	Rubisco cycle	This study
Pr-RBCL-R	CCGTAGGTGCCHGCCTTCCACA	64.6	*Prorocentrum sp.*	Rubisco cycle	This study
Pr-LH-F1	CAGGCCTGAAGTTCTCGGAC	57.8	*Prorocentrum sp.*	Light-harvesting center	This study
Pr-LH-R1	TCAGGATCCGAGGACGTGATC	58.2	*Prorocentrum sp.*	Light-harvesting center	This study
Pr-LH-F2	TCCGACATCTTCATCAAGCCTC	56.7	*Prorocentrum sp.*	Light-harvesting center	This study
Pr-LH-R2	ATGCTGCCGATGGCCTTGTGGTG	64.6	*Prorocentrum sp.*	Light-harvesting center	This study
Pr-PSB-F	ATACAATGGAGGAACATATCAATT	50.2	*Prorocentrum sp.*	psbA protein	This study
Pr-PSB-R	ACWAYGAATACTGCTGATGC	51.5	*Prorocentrum sp.*	psbA protein	This study
Ch-RH-F	TTGGCTGTTAACCGTGCCAC	58.5	*Chaetoceros socialis*	Rhodopsin	This study
Ch-RH-R	AACGACCACCTAGATTGTCTT	53.5	*Chaetoceros socialis*	Rhodopsin	This study
Ch-RBCL-F	GCTTTAGAAGCAATGGTTATGGCT	55.8	*Chaetoceros socialis*	Rubisco cycle	This study
Ch-RBCL-R	ATTTAGCAGCATCACGTAAGATTAC	53.4	*Chaetoceros socialis*	Rubisco cycle	This study
Ch-LH-F	GTACAGTACTGTGCCTGGTTCGC	59.8	*Chaetoceros socialis*	Light-harvesting center	This study
Ch-LH-R	GGTGCTTCATTACACCACAGCG	58.7	*Chaetoceros socialis*	Light-harvesting center	This study
Ch-PSB-F	TAGCAGCGGCTTCTGCTGTA	58.8	*Chaetoceros socialis*	psbA protein	This study
Ch-PSB-R	CGGATTAAAGATGAAGTTACTA	47.5	*Chaetoceros socialis*	psbA protein	This study
Ph-RH-F	TACCACTACCTCCGCATCTTCA	57.4	*Phaeocystis pouchetii*	Rhodopsin	This study
PH-RH-R	AGCAGAAGATGATCTCGATGAG	54.0	*Phaeocystis pouchetii*	Rhodopsin	This study
Ph-PSB-F	AGCTGCTTCAATCGAYGART	54.5	*Phaeocystis pouchetii*	psbA protein	This study
PH-PSB-R	GACCGATTGGGTAGATTACG	52.5	*Phaeocystis pouchetii*	psbA protein	This study
PN-RH-F	ACTCGCTGGGTCTACTGGAT	57.6	*Pseudo Nitchzia*	Rhodopsin	This study
PN-RH-R	CAGGGTAGGTGCACCAAGAG	57.6	*Pseudo Nitchzia*	Rhodopsin	This study
PN-RBCL-F	TGCTGCAGTAGCTGGTGAATCTT	58.8	*Pseudo Nitchzia*	Rubisco cycle	This study
PN-RBCL-R	AAAGGRAAGAATACATCTGTTGT	51.2	*Pseudo Nitchzia*	Rubisco cycle	This study
PN-PSB-F	TTAGGTGCTCACGTAGCACA	56.1	*Pseudo Nitchzia*	psbA protein	This study
PN-PSB-R	TACGGTACAGTTGTAGTGAT	50.1	*Pseudo Nitchzia*	psbA protein	This study

Table S4. 3rd set of primers designed and ordered from literature, targeting photosynthesis machinery and rhodopsin. Species specific for *Karlodinium veneficum, Prorocentrum* sp., *Chaetoceros socialis, Phaeocystis* sp., and *Pseudo Nitzschia* sp..

Primer name	Sequences (5'-3')	Tm °C	Target organism	Target gene	Reference
KarloF1	ATGGCTGGATGGTTCTTCATC	55	*Karlodinium Veneficum*	Rhodopsin	This study
KarloR1	CAACKGTGACGATGAAACGC	55.8	*Karlodinium Veneficum*	Rhodopsin	This study
KarloF2	CCATCGTCTATCGTTACAT	48.8	*Karlodinium Veneficum*	Rhodopsin	This study
KarloR2	ATCARRCGCCAGAACATGGC	58.5	*Karlodinium Veneficum*	Rhodopsin	This study
KarloF3	GTTCACTACTTCTACATGCGG	53.1	*Karlodinium Veneficum*	Rhodopsin	This study
KarloR3	TAGARCTCGATCATCTGCA	51.5	*Karlodinium Veneficum*	Rhodopsin	This study
KarloF4	TGCAGATGATCGAGYTCTA	51.5	*Karlodinium Veneficum*	Rhodopsin	This study
KarloR4	CACCCGCGAARATCTCGA	56.2	*Karlodinium Veneficum*	Rhodopsin	This study
KarloF5	CTTCATCYTSTTCGAGRT	48.5	*Karlodinium Veneficum*	Rhodopsin	This study
KarloR5	GTGACRATGAANCGCATG	51.8	*Karlodinium Veneficum*	Rhodopsin	This study
KarloF6	GTTCCNYTGCAGATGATC	50.6	*Karlodinium Veneficum*	Rhodopsin	This study
KarloR6	GTGACRATGAACCGCATGGT	57.0	*Karlodinium Veneficum*	Rhodopsin	This study
ProroF1	AGATAACGACGCATACCG	51.7	*Prorocentrum sp.*	Rhodopsin	This study
ProroR1	GTAGCCAGAGATGATCAT	48.1	*Prorocentrum sp.*	Rhodopsin	This study
ProroF2	GCGTACCGTTACATGGATT	52.5	*Prorocentrum sp.*	Rhodopsin	This study
ProroR2	GTAACCAGAGATGATCAT	45.4	*Prorocentrum sp.*	Rhodopsin	This study
DinorhodF2N	CTCTCGTTCGGCATHGCNGCNATGGG	65.8	Dinoflagellates	Rhodopsin	Huang et al. 2019
DinorhodR3N	CTCACCNGGRTASCCCARNGCNACCAT	66.4	Dinoflagellates	Rhodopsin	Huang et al. 2019
UncultF1	TACATGTACATGCGTGAGTA	50.7	General Eukaryote	Rhodopsin	This study
UncultR1	TCAATCATCTGCAGAGGGAC	54.3	General Eukaryote	Rhodopsin	This study
UncultF2	TCATTCTGTTCGAAATCTTC	47.8	General Eukaryote	Rhodopsin	This study
UncultR2	AGTGGGTAGATGGACCAGCC	58.9	General Eukaryote	Rhodopsin	This study
KarlorhodF1	GATCCTCTCTGCRGTGCAGC	59.4	*Karlodinium Veneficum*	Rhodopsin	This study
KarlorhodR1	ACGAACAAGATGAAGAACCA	51.7	*Karlodinium Veneficum*	Rhodopsin	This study
KarlorhodR2	CTAGAAAGATARCCRACAGT	48.9	*Karlodinium Veneficum*	Rhodopsin	This study
KarlorhodF2	TCCTCTCTGCAGTGCAGC	57.5	*Karlodinium Veneficum*	Rhodopsin	This study
KarlorhodR3	TGAACCGCATGGTTGAAAAA	53.4	*Karlodinium Veneficum*	Rhodopsin	This study
KarlorhodR4	CCTTCRCCCGCGAARATCTC	58.4	*Karlodinium Veneficum*	Rhodopsin	This study

Table S5. 4th set of primers designed and ordered from literature.

Primer name	Sequence (5'-3')	Tm°C	Target organism	Target area	Reference
OP41	CGACGCATGCACGCATTTTTTTTTTTTTTT	59	Eukaryotes	poly-A tail	Vader et al., 1999
OP283	AAGCGACGCATGCACGCATT	62	Eukaryotes	poly-A tail	Vader et al., 1999

Table S6. Poly-A primers.

PCR trials

Primer pairs, Annealing and melting temperatures, temperatures at which the primers were tested, and whether or not they were successful. 1st temperature: annealing temperature recommended by NEB®/ 2nd temperature: PCR temperature. Many of the PCRs were repeated for optimization, therefore these tables are not quantitatve.

Results	No bands/ primer dimers	Some bands	Strong bands of right size
Standard PCR			
Temperature trial results			

Table S7. Legend for PCR trial figures indicating: 1. Negative result (red), 2. Some bands (yellow), 3. Positive result (green), temperature trial results indicated by striped cells.

Target group	Primer	proteo R1	proteo R2	uncult R1	uncult R2	uncult R3	xantho R1	xantho R2	xantho R3	xantho R4	phaeo R1	phaeo R2	rbcl-A (R)	rbcl- B (R)	Lhcf (R)	psbAK (R)	Diatom18sRL	Dino18sRL	RbclR	RhBR1	RhBR2	RhBR3	Dino GR1	Dino GR2	18S V4 R	1209 R	poly-A OP283
Blue Proteorhodopsin	proteo F1	48/48 & 50																									
	proteo F2		51/50 & 49-55																								
	blueF	49/ 50	51/ 50 & 53																	46/ 50	47/ 50						53
	dino-SL	49/ 50	51/ 50 & 53				59/ 51-58	50/ 51-58	54/ 51-58	50/ 51-58										46/ 50	47/ 50		46/ 50	49/ 50			53
	uncult F1			54/ 53	54/ 53-55	57/ 53-55																					
	uncult F2			51/ 51	51/ 51	51/ 51																					
	uncult F3			54/ 53	54/ 53	57/ 53																					
Green xanthrorhodpsin Prorocentrum	xantho F1								48/ 49	48/ 49																	
	xantho F2						47/ 53	47/ 53																			
	green F1						50/ 49-50	50/ 49 & 53-55																			
	green F2						54/ 51	50/ 51																			
Green xanthrorhodpsin Phaeo/ Pseudo.	phaeo F1										49/ 51 & 4	48/ 51															
Regulation of Rubisco activity- halotype A	rbcl- A(F)												48/ 48														
Regulation of Rubisco activity- halotype B	rbcl-B(F)													48/ 48													
Light harvesting complex	Lhcf (F)														56/ 49-55												
psbA gene	psbAK (F)															53/ 50-51											
General primers	Euk528F																?/ 51 &53	51-53								53/ 50	
Rubisco, 2nd attempt	RbclF595																		57/ 53								
Rhodopsin sequence	RhbF1																			50/ 50	51/ 50						53
	RhbF2																			50/ 50	51/ 50						53
	RhbF3																					48/ 50					
Dinoflagellate	Dino GF1																						50/ 50	53/ 50			53
	Dino GF2																						50/ 50	53/ 50			53

Table S8. PCR trial results with 1st and 2nd set of primers.

Species specific primers	Primer	K-RH-R	K-RBCL-R	K-LH-R	K-PSB-R	Pr-RH-R1	Pr-RH-R2	Pr-RBCL-R	Pr-LH-R1	Pr-LH-R2	Pr-PSB-R	Ch-RH-R	Ch-RBCL-R	Ch-LH-R	Ch-PSB-R	Ph-RH-R	Ph-PSB-R	PN-RH-R	PN-RBCL-R	PN-PSB-R
	K-RH-F	58/ 56																		
	K-RBCL-F		52/ 50																	
	K-LH-F			56/ 56																
Karlodinium verificum	K-PSB-F				53/ 50															
	Pr-RH-F					54/ 50														
	Pr-RH-F2						44/ 50													
	Pr-RBCL-F							67/ 56												
	Pr-LH-F1								61/ 56											
	Pr-LH-F2									58/ 56										
Prorocentrum	Pr-PSB-F										50/ 50									
	Ch-RH-F											55/ 50 & 56								
	Ch-RBCL-F												53/ 50							
	Ch-LH-F													61/ 56						
Chaetoceros socialis	Ch-PSB-F														47/ 50					
	Ph-RH-F															55/ 50				
Phaeocystis	Ph-PSB-F																54/ 50			
	PN-RH-F																	60/ 56		
	PN-RBCL-F																		50/ 50	
Pseudo Nitchzia	PN-PSB-F																			51/ 50

Table S9. PCR trial results with species-specific primers.

	Primer name	KarloR1	KarloR2	KarloR3	KarloR4	KarloR5	KarloR6	ProroR1	ProroR2	DinorhodR3N	UncultR1	UncultR2	KarlorhodR1	KarlorhodR2 (vector)	KarlorhodR3	KarlorhodR4
Karlodinium verificum	KarloF1	50														
	KarloF2		50													
	KarloF3			50												
	KarloF4															
	KarloF5					50										
	KarloF6						50									
Prorocentrum sp.	ProroF1															
	ProroF2								45 & 50							
Dinoflagellate	DinorhodF2N									50						
Uncultured microorganism	UncultF1										45 & 50					
	UncultF2											45				
Karlodinium verificum	KarlorhodF1					45							45 & 50		45	
	KarlorhodF2													50	45 & 50	

Table S10. PCR trial results with final primers designed for *Karlodinium* veneficum and *Prorocentrum* sp. for qPCR assay

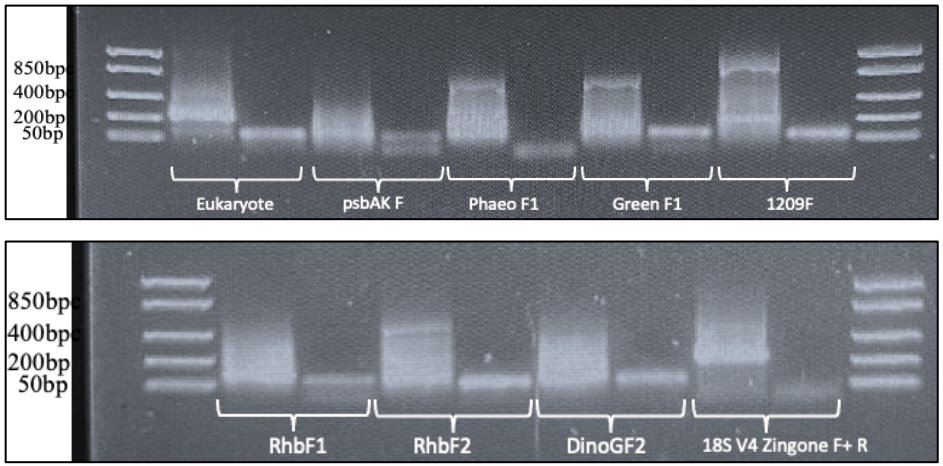


Figure S3. Results on 2% Agarose gel using 3µl GeneRuler Low Range Ladder from Thermo Scientific™ showing PCR products at 50°C with forward primers and OP41 reverse primer on poly-A cDNA from polar night. Results showed large, unspecific products, likely due to the use of OP41.

Sequence alignments for qPCR assay primers

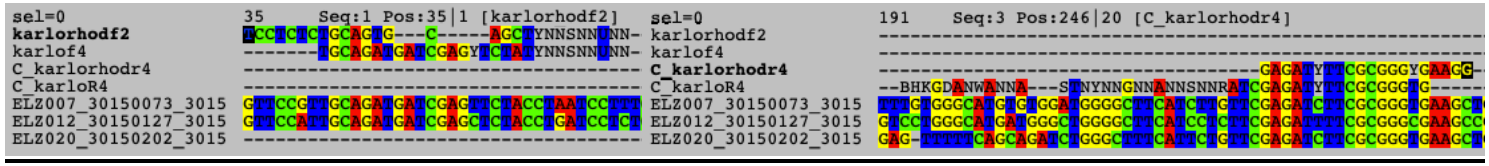


Figure S4.A. Multiple Sequence alignment for *Karlodinium veneficum* qPCR assay primer pairs KarloF4/R4 and KarlorhodF2/R4 using Seaview.exe, aligned with Sanger sequencing results which showed a Hit on *K. veneficum* rhodopsin. The alignment shows the sequence overlap between the forward and reverse primers, respectively. Position within aligned sequences indicates expected amplicon sizes to be: KarloF4/R4= ~200bp and KarlorhodF2/R4= ~211bp.

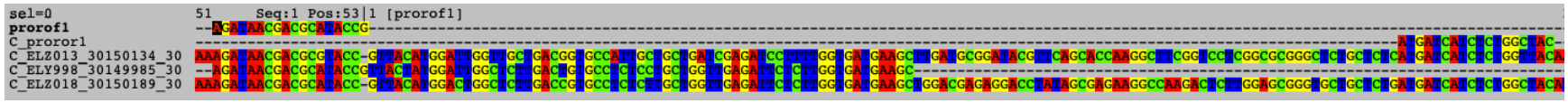


Figure S4.B. Multiple Sequence alignment for *Prorocentrum* sp. qPCR assay primer pair ProroF1/R1 using Seaview.exe, aligned with Sanger sequencing results which showed a Hit on *Prorocentrum donghaiense* rhodopsin. Position within aligned sequences indicates expected amplicon size to be ~154bp.

qPCR

Standard Curve Amplification plots

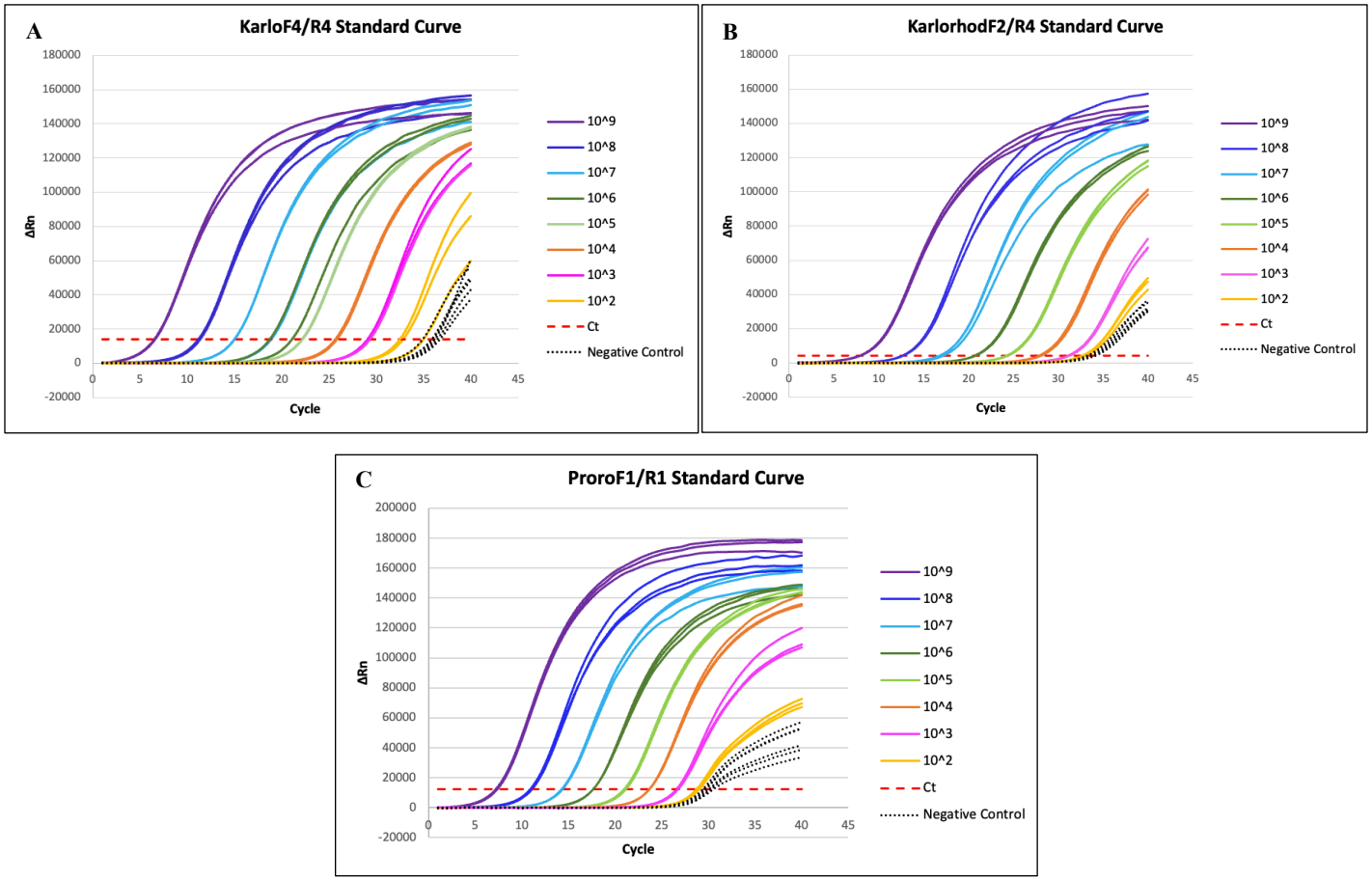
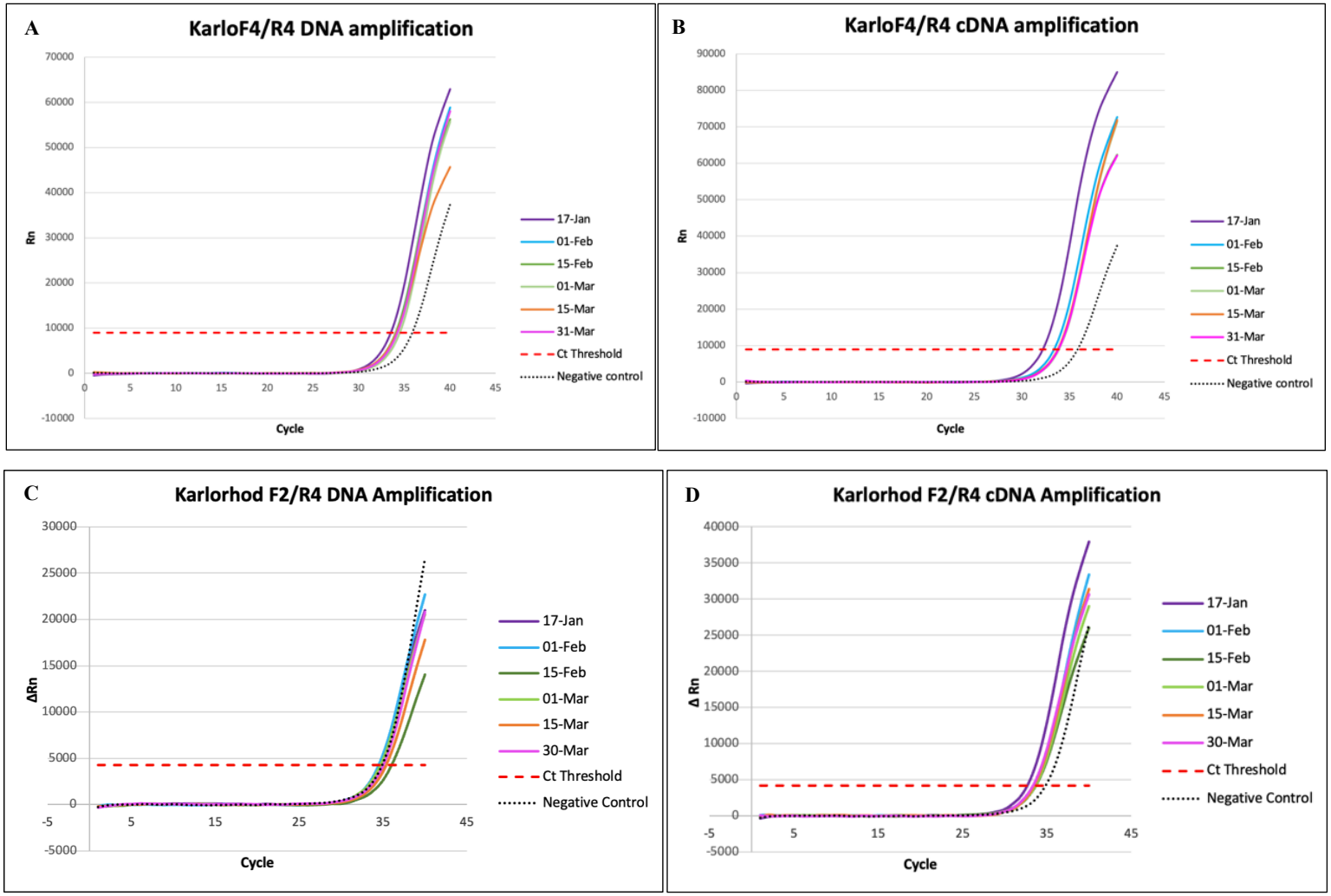


Figure S5 (A, B, C). Standard Curve amplification plots for primer pairs used in qPCR assay.

Assay Amplifcation plots



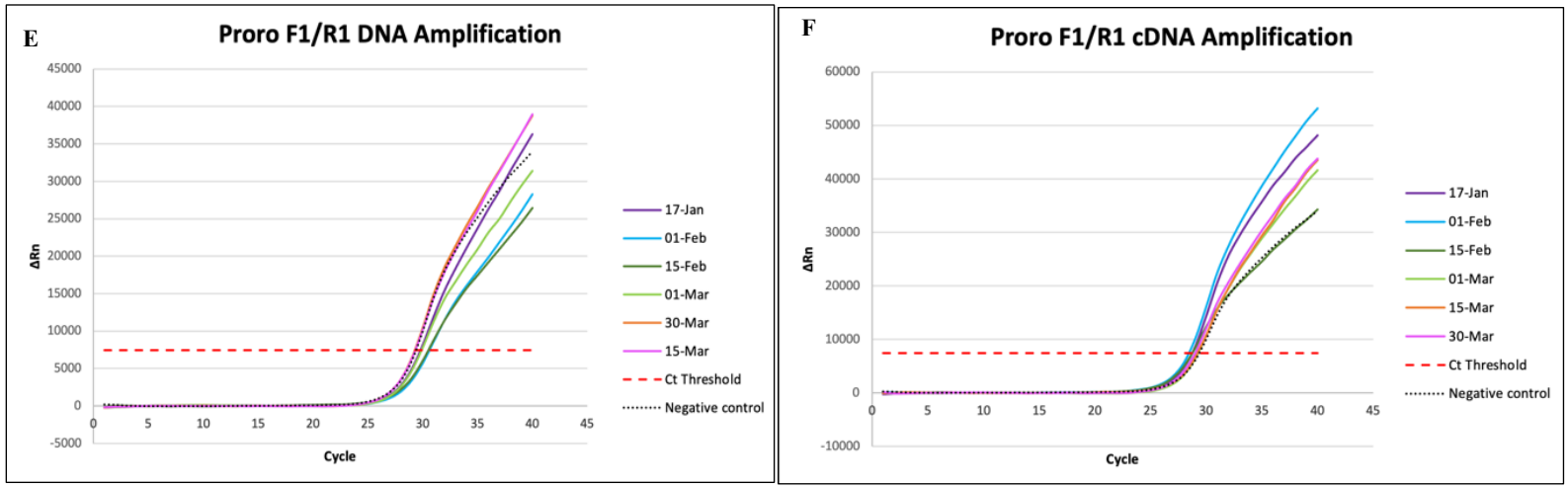


Figure S6 (A, B, C, D, E, F). Amplification plots for qPCR Assays comparing DNA vs. cDNA amplification for each primer throughout time series. 1. KarloF4/R4: Ct= 8996,918 2. KarlorhodF2/R4: Ct= 4231.96

Assay Melt Curves

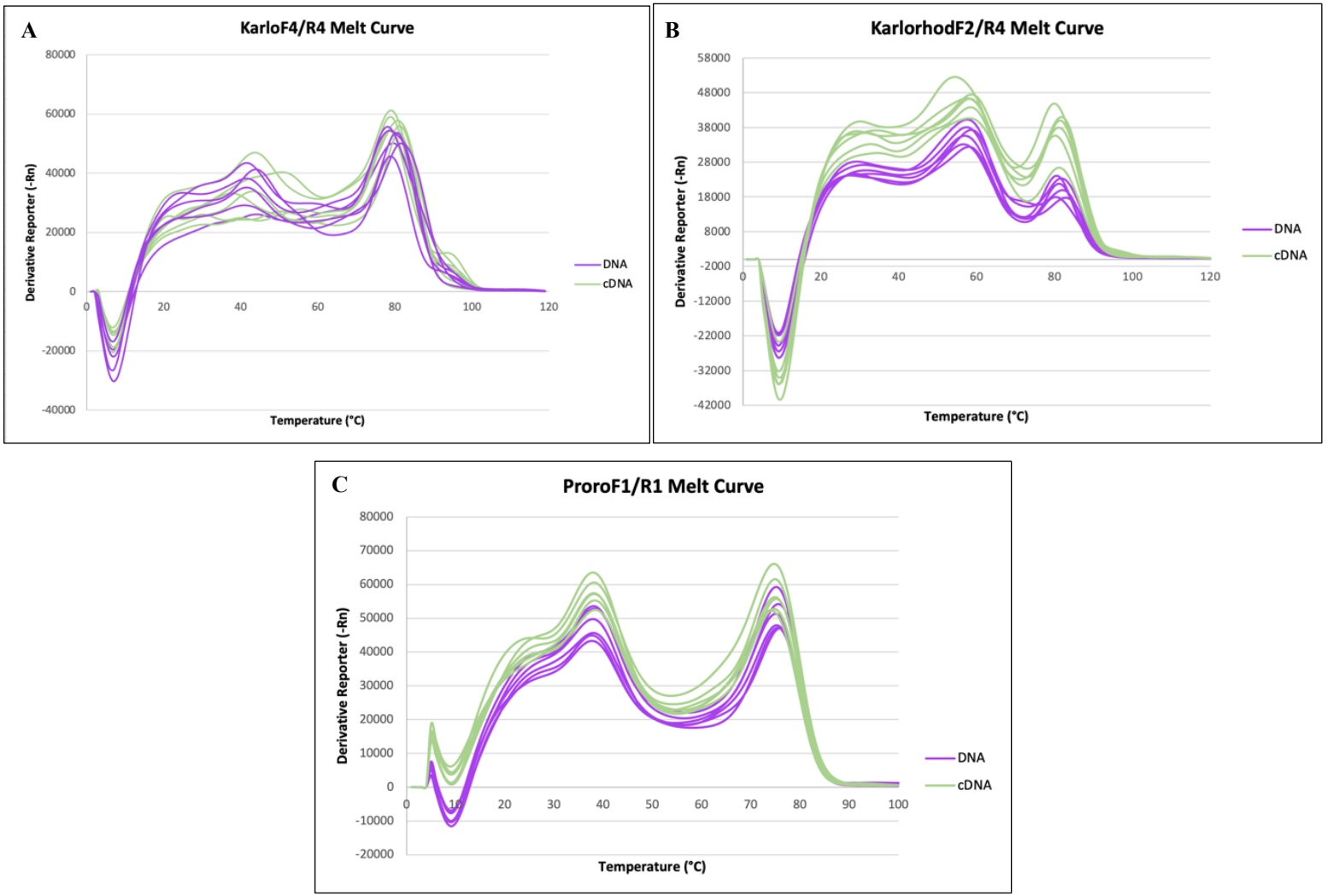


Figure S7 (A, B, C). Melt Curves for qPCR assays comparing DNA and cDNA T_ms in each primer pair.

Copy numbers

Table S11. Copy numbers *(mL^{-1})* for each qPCR primer pair.

	KarloF4/R4		KarlorhodF2/R4		ProroF1/R1	
Date	**DNA**	**RNA**	**DNA**	**RNA**	**DNA**	**RNA**
17/01/23	17,4972234	155,240776	29,8819646	471,841825	12,4325349	101,526991
01/02/23	12,6688636	72,4138591	27,9107647	494,387335	12,0722682	147,112378
15/02/23	10,0277562	44,5996637	17,9459285	308,786368	7,32842376	102,465183
01/03/23	7,33396004	38,0028382	26,5463206	293,336581	10,9602852	73,5666991
15/03/23	8,32074007	58,5685752	25,4911186	499,779967	18,064082	98,6644344
31/03/23	8,85759949	46,7389159	25,028011	429,164335	12,3488698	117,500399

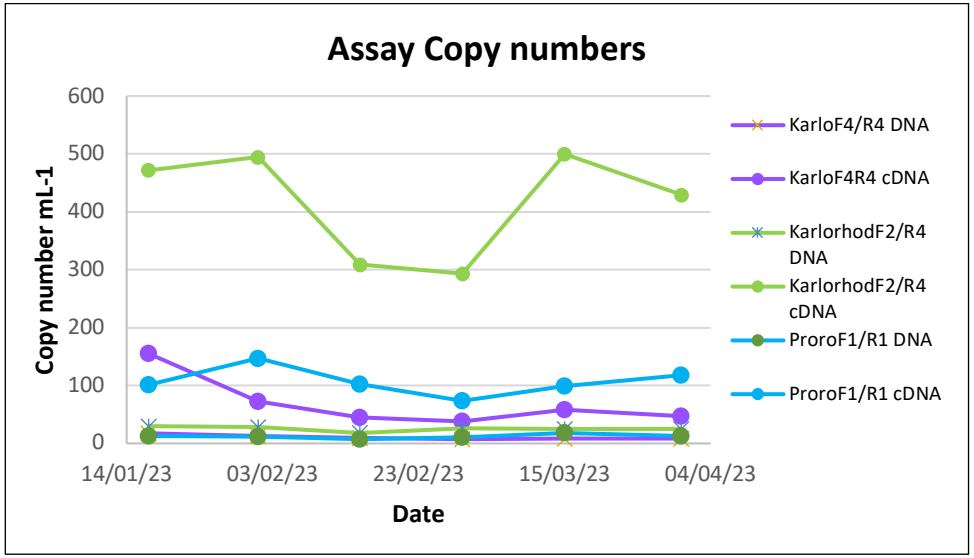


Figure S8. DNA and cDNA copy numbers (mL^{-1}) from each assay. Purple: KarloF4/R4, Green: KarlorhodF2/R4 and Blue: ProroF1/R1.

Normalized environmental copy numbers (mL^{-1}) from qPCR Assays

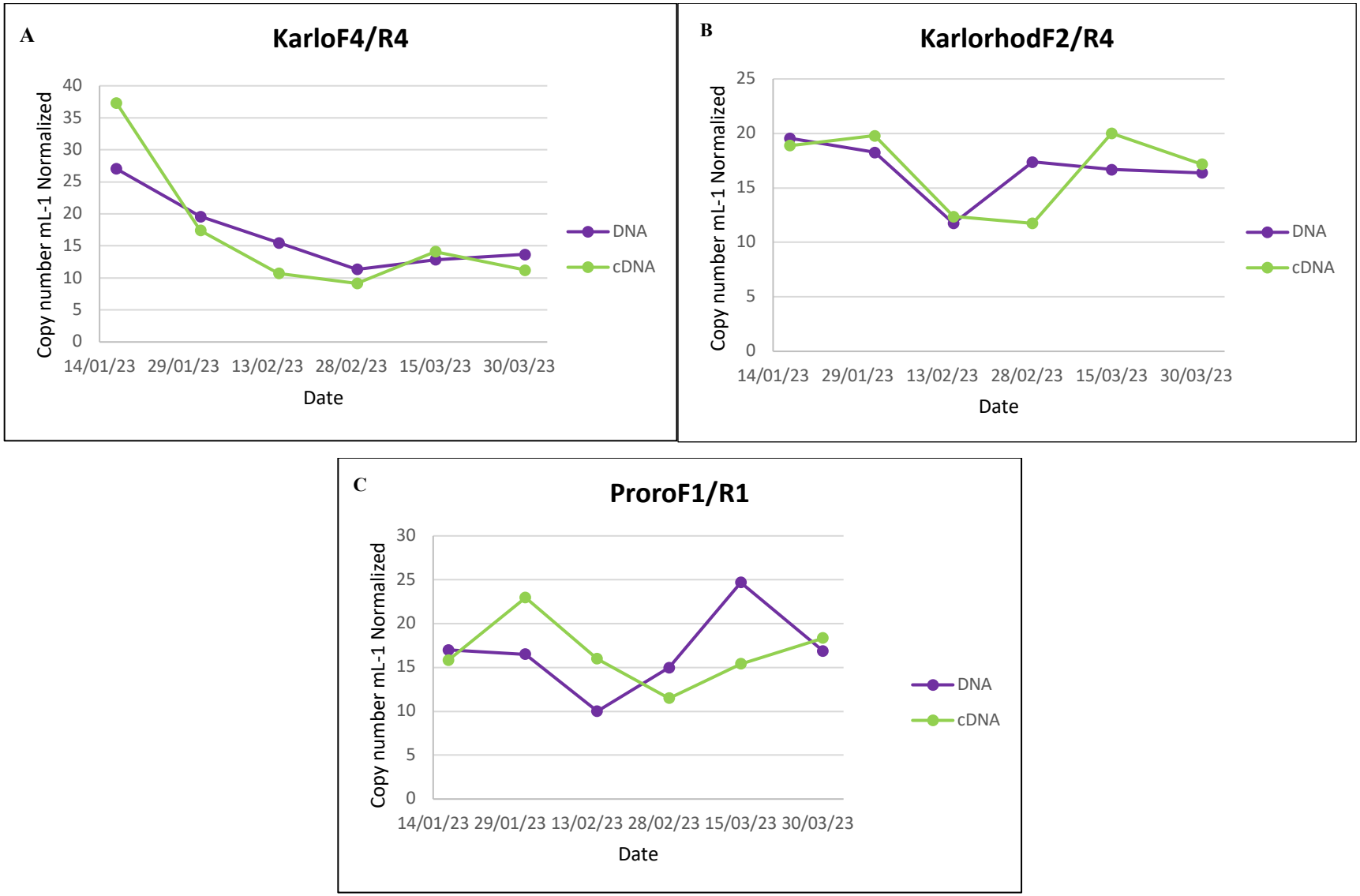


Figure S9. Copy numbers/ mL of filtered seawater of target DNA and cDNA in environmental samples, normalized for comparison.