An eDNA assessment of pelagic community composition in Northern California

Parker K Lund, University of British Columbia

Mentors: Jacoby Baker, Dr. Francisco P. Chavez, Dr. Kathleen Pitz

Summer 2023

Keywords: eDNA, Northern California, community ecology

ABSTRACT

The California Current System is one of four eastern boundary upwelling systems, which are among the most productive marine ecosystems in the world. While several long-term biomonitoring programs have been established throughout the ecosystem, most of these efforts target specific taxa that are either ecologically important or economically valuable. Here, we captured a broad diversity of organisms using environmental DNA (eDNA) as a relatively low-cost method for tracking variation in pelagic community composition. In northern California, community composition is primarily driven by depth and season. The major taxa driving variation include hydrozoa, haptophytes, chlorophytes, diatoms, and other algal groups. In the surface community, the copepod *Pseudocalanus mimus* was another key player, making up a large percentage of reads throughout the spring and summer upwelling season. Different diatom classes displayed their own seasonal patterns, though Coscinodiscophyceae was in the highest relative abundance, starting in late winter and peaking in spring. Surprisingly, a high percentage of dinoflagellate DNA was detected throughout most of the year, although this was a particularly strong year for upwelling that may not reflect average conditions. Long-term biodiversity monitoring programs are critical not only for tracking how ecosystems are changing in response to

climate change, but also for the management of natural resources and marine protected areas.

INTRODUCTION

As ocean warming continues to accelerate (Cheng et al., 2022), it is vital to establish continual biodiversity monitoring programs that are sustainable and cost-effective. Many traditional monitoring methods are labor intensive (Harvey et al., 2017), require significant expertise to identify organisms that are difficult to distinguish morphologically (Bailey et al., 2016), and often result in missing species that are less visually obvious (Lamy et al., 2021). One alternative method is using environmental DNA (eDNA) collected on filters from water, soil, or air samples, to identify a large array of organisms through a process called metabarcoding. In marine samples, eDNA from larger organisms is extracted from tissue, skin, scales, and fecal matter that have been shed into the water column. For smaller organisms like plankton and bacteria, eDNA is often extracted from the whole organism and living cells. Metabarcoding allows us to identify different taxa in a mixed sample by comparing short, targeted gene sequences, or "barcodes", to known sequences in a database. The gene targets are both hypervariable and highly conserved, meaning they have enough variation to distinguish different taxa without mutating too quickly over time (Chavez et al., 2021). While we are currently unable to estimate the absolute abundance of different organisms using metabarcoding alone, we can determine broadly which organisms are present and how environmental variables impact community composition on a large scale.

The Marine Biodiversity Observation Network (MBON) was established to coordinate global monitoring efforts (Muller-Karger et al., 2014), including the well-studied California Current System along the west coast of North America. The California Current System is one of four eastern boundary upwelling systems, which are among the most productive marine ecosystems in the world (Mackas et al., 2006). It is characterized by strong upwelling in the spring and summer, particularly between the 34°N and 44°N latitudes, where wind-driven Ekman transport pushes coastal surface waters offshore, drawing up cold nutrient-rich water from around 60 m deep (Chavez & Messié, 2009).

Although seasonal cycles of upwelling and primary production generally align (Chavez & Messié, 2009), latitudes with the greatest water volume transport do not necessarily display the highest chlorophyll concentrations. Southern latitudes typically experience greater overall water transport, however this is coupled with a lower average chlorophyll concentration than northern latitudes. The especially high productivity in the northern part of the system is likely enhanced by geographic features, such as a wide continental shelf and an influx of nutrients from large freshwater sources like the Columbia River, which help sustain higher chlorophyll concentrations throughout the year (Hickey & Banas, 2008).

Here, we present a preliminary eDNA analysis of the newly sampled Trinidad Head Line in northern California, which lies in a transition zone between the two major portions of the California Current System (Robertson & Bjorkstedt, 2020). We used the COI and 18S gene markers to target invertebrates and plankton, identifying how the community dynamics shift across space and time and determining which taxa are most responsible for variability. For both gene targets, differences in community composition were predominately driven by depth, followed by season, with distinct taxa occurring particularly towards the surface and during the wintertime. These efforts enhance the existing monthly ecosystem monitoring that has been conducted along the Trinidad Head Line since 2006 (Bjorkstedt & Peterson, 2015) as part of the larger Marine Biodiversity Observation Network operating throughout the California Current System (Muller-Karger et al., 2014).

MATERIALS AND METHODS

SITES

The Trinidad Head Line (THL) lies along the 41°N latitude in northern California, within an embayment located between two major upwelling centers: Cape Blanco and Cape Mendocino (Robertson & Bjorkstedt, 2020). North of Cape Blanco, upwelled waters predominately stay nearshore, however in this region sweeping jets of cold water often extend far past the shelf break (Barth et al., 2000). Although substantial water transport occurs at this latitude, there is notable interannual variability in chlorophyll

concentration (Hickey & Banas, 2008). The THL is comprised of five stations, three of which are located nearshore over the shelf and two of which extend beyond the shelf break that lies 18 km offshore at a depth of 200 m (Robertson & Bjorkstedt, 2020).

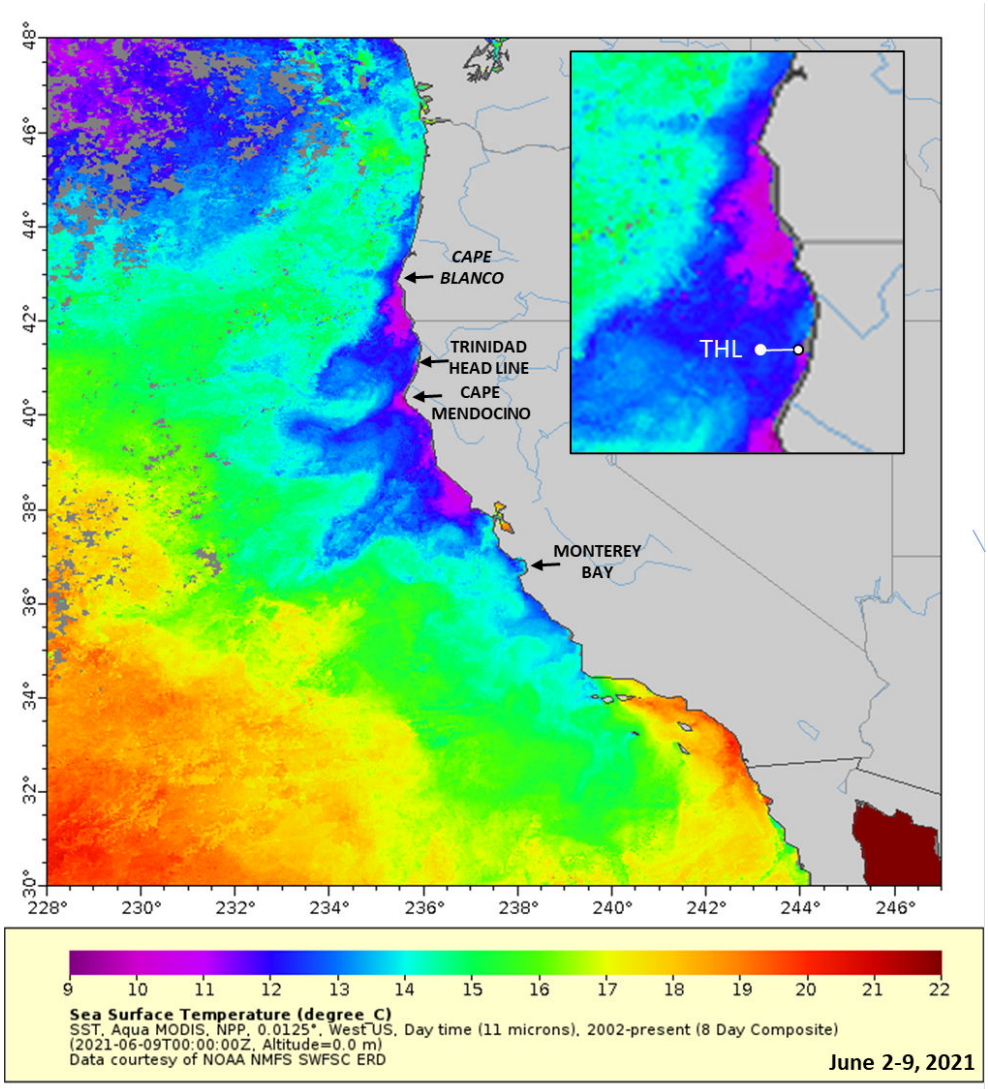

Figure 1. Map of sea surface temperatures along the west coast of the United States of America. Temperatures (in Celsius) represent an average between June 2nd through June 9th, 2021, during the summer upwelling period. Major landmarks around the Trinidad Head Line have been labeled, including the two largest upwelling centers. In the upper right, a zoomed-in image displays the location and path of the Trinidad Head Line (THL). The small white dot on the right is station TH01; the large white dot on the left is station TH05.

SAMPLING PROCEDURE

Sampling was conducted monthly (except August and October) over the course of a 12-hour cruise from June 22nd 2021 to May 18th 2022, with the nearshore stations sampled in the late afternoon and the offshore stations sampled after dark. This split sampling schedule reduced navigational difficulties, while still capturing vertical diel migration in the offshore stations. Hydrographic data was measured using a Sea-Bird Electronics 19plus V2 SEACAT profiler CTD, integrated with the Sea-Bird Electronics ECO-55 water sampler affixed with Niskin bottles. At each station, two to five different depths were sampled depending on the bottom depth (TH01: 35 m; TH02: 75 m; TH03: 140 m; TH04: 460 m; TH05: 780 m bottom depth). One liter of water was collected from each Niskin bottle in sterile Whirl-Pak bags, which were immediately placed on ice for the duration of the cruise. The samples were stored overnight in coolers with fresh ice, then filtered within 24 hours through a 0.22 um Poretics filter using a peristaltic pump. The pumps were allowed to continue running until all excess water was removed and the filters appeared dry. To control for contamination, one liter of sterile water per cruise was filtered at the same time using the same process and materials. Tweezers for transferring the filters were sterilized in the field with RNase Away, rinsed twice with Milli-Q water, and once in ethanol. The filters were transferred into cryovials and frozen at -80°C until they were shipped to the Montrey Bay Aquarium Research Institute for DNA extraction.

SEQUENCING

Environmental DNA was extracted using the Qiagen DNeasy 96-well Blood and Tissue Kit, which included a blank extraction of sterile water to serve as a negative control for each plate. PCR was used to amplify three target amplicons: cytochrome c oxidase subunit I (COI), 18S rRNA (region v9), and 12S rRNA. Sterile water was used again as an additional negative control during PCR amplification. Although only the COI and 18S genes were analyzed for this project, the 12S gene was still sequenced and will be analyzed in the future. PCR products were purified using the Agencourt AMPure XP bead system, followed by gel electrophoresis to verify only the target amplicons were retained. DNA

concentration was measured using fluorometry before shipping the PCR products to Michigan State University's Genomics Core for sequencing via Illumina MiSeq.

Table 1. Overview of gene targets, primers, and thermalcycler conditions for PCR amplification. Includes the expected length of the fragments (in base pairs), as well as the literature reference for each primer set.

Gene	Taxa	Primers	Primer Sequences and PCR Conditions	Reference
COI	Invertebrates	mlCO1intF HCO2198 313bp	5'-GGWACWGGWTGAACWGTWTAYCCYCC-3' 5'-TAAACTTCAGGGTGACCAAAAAATCA-3' 95 °C for 10 minutes 16 cycles of: • 94 °C for 10 seconds • 62 °C for 30 seconds (this changes -1°C for each subsequent cycle) • 68 °C for 60 seconds Then 25 cycles of: • 94 °C for 10 seconds • 46 °C for 30 seconds • 68 °C for 60 seconds A final elongation step of 72 °C for 10 minutes Hold at 4 °C	Leray et al. 2013 Folmer et al. 1994
18S	Phytoplankton	1391F EukBr ~260 bp	5'-GTACACACCGCCCGTC-3' 5'-TGATCCTTCTGCAGGTTCACCTAC-3' 95° C for 10 minutes 35 cycles of: • 94° C for 45 seconds • 57° C for 30 seconds • 68 °C for 90 seconds Final elongation step of 72° C for 10 minutes Hold at 4 °C	Amaral-Zettler et al. 2009 Stoek et al. 2010 Earth Microbiome Project (EMP)
12S	Vertebrates	MiFish-U-F MiFish-U-R 163-185 bp	5'-GTCGGTAAAACTCGTGCCAGC-3' 5'- CATAGTGGGGTATCTAATCCCAGTTTG-3' 98°C 30 seconds 38 cycles of: • 98°C 10 seconds • 62°C 10 seconds • 72°C 30 seconds A final elongation step of 72°C for 5 minutes Hold at 4°C	Miya et al. 2015

UPWELLING INDICES

Upwelling was examined using two indices: the Coastal Upwelling Transport Index (CUTI) and the Biologically Effective Upwelling Transport Index (BEUTI). CUTI estimates the amount of water being vertically transported in and out of the surface mixed

layer, while BEUTI converts that estimate to the quantity of nutrients being upwelled using the concentration of nitrate at the bottom of the mixed layer (Jacox et al., 2018). Upwelling indices were binned by latitude (41°N latitude) and covered 30 km of coastline. A running average was calculated for each day, where the indices were averaged over five days (including two days before and two after the cruise date).

BIOINFORMATICS

The COI and 18S rRNA sequence data was processed as described in (Min et al., 2023) using Atropos (Didion et al., 2017) to remove primer sequences and DADA2 (Callahan et al., 2016) to trim regions of poor quality, filter out chimeric sequences, and merge paired forward/reverse reads. After merging, taxa was assigned by searching the NCBI GenBank database with the blastn algorithm (NCBI Resource Coordinators, 2018), in conjunction with MEGAN6's lowest common ancestor algorithm (Huson et al., 2016). Phyloseq (McMurdie & Holmes, 2013) was used to remove non-target taxa from the dataset, including bacteria and terrestrial vertebrates. There were many COI amplicon sequence variants (ASVs) miscategorized as terrestrial arthropods, including a large group of terrestrial spiders in the order Araneae. This group of ASVs was excluded because it placed Araneae among the top 20 most abundant taxonomic orders, however their exclusion did not affect the general patterns seen in beta diversity. ASVs that were abundant in the negative controls were also removed, however none of them were highly prevalent in the regular samples. Four samples were also removed because they clustered with communities at other depths and it appeared that the recoded ocean depths had been swapped between samples, however these should be looked at more closely to verify that their removal was appropriate.

Differences in beta diversity between categories (station, depth, season) were assessed in QIIME2 (Bolyen et al., 2019) with the DEICODE plugin (Martino et al., 2019) using a robust Aitchison PCA (RPCA) and an analysis of similarities (ANOSIM) test with 999 permutations. Datasets were not normalized beforehand, as the RPCA is designed to handle community ecology data that often displays a large percentage of counts that are zero. Depth and season bins were generated based on the natural clustering of the samples

in PCA space and when plotting them by PC1 and PC2. Binning was performed for the COI PCA results and may need to be adjusted later for 18S, since different taxa are targeted by each gene marker. To identify top ASVs driving differences in PC1 and PC2, the loading scores for all ASVs were extracted from the RPCA and the relative abundance of ASVs with the largest positive and negative scores were plotted against the environmental variable associated with a given PC. ASVs with large loading scores that were not assigned taxonomy be MEGAN6 were checked manually against the NCBI GenBank database and assigned a putative identity typically at the phylum level. Any unidentified ASVs that were likely bacteria were not plotted (this was only common in PC1 for the COI gene). Differences in beta diversity were also separately assessed for the surface layer (samples collected < 25 m deep), but only for the COI gene due to time constraints. Throughout the following figures, the relative abundances of different taxa are plotted with different scales for the y-axis because there is such a wide range of values (i.e. some organisms comprise up to 30% of the reads per sample, while others may only account for 0.05%). Log transformations were attempted, however plotting different organisms along the same axis ultimately obscured changes in pattern of relative abundance across time and space.

RESULTS

UPWELLING AROUND THE TRINIDAD HEAD LINE

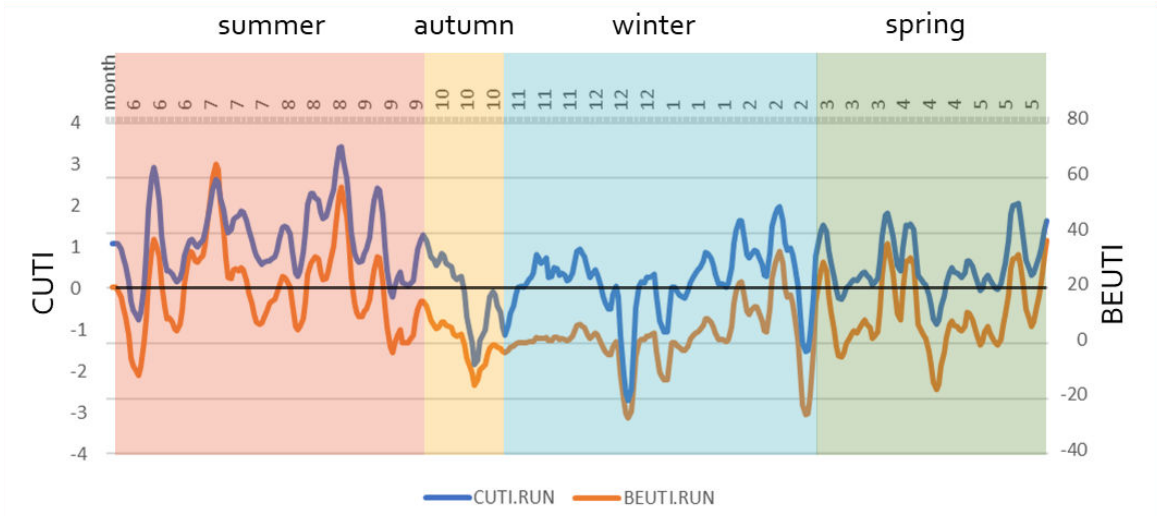

Figure 2. The five-day running average of the CUTI (m^2/s) and BEUTI (umol/L) indices for the 41°N latitude from June 2021 through May 2022. The black line highlights the time periods highlights when water transport is zero for the CUTI index. The season bins are defined by how the COI data clusters between months, making it difficult to bin August and October since those months were not sampled.

There was a positive linear correlation between the daily CUTI and BEUTI indices for June 2021 – May 2022 ($df = 363$, $R2 = 0.88$, $P < 0.001$), although nitrate concentrations rarely dipped below zero even in the winter months (Figure 2). The greatest upwelling occurred during the summer months, particularly July-August, in contrast to the colder months of October-January where the CUTI running average remained negative. Notably, 2021 represents a particularly anomalous year for both cumulative volume transport and upwelled nitrate in comparison with upwelling from the past 35 years (Figure 3). Although 2022 initially starts strong, it ends up being a fairly average year for cumulative transport.

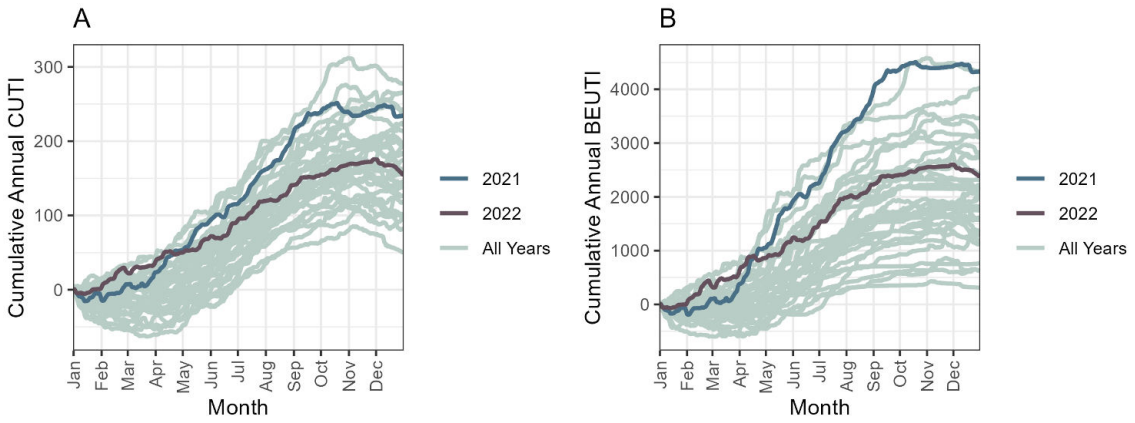

Figure 3. Cumulative annual upwelling from 1988-2023 for the 41°N latitude of the California Current System. Both indices are calculated daily, with the (A) CUTI index estimating the volume of water transported in m^2/s and the (B) BEUTI index estimating the concentration of upwelled nitrate in umol/L.

COMMUNITY DYNAMICS FOR THE COI GENE

For the COI gene, the community composition was significantly different between depth bins (R = 0.490, P = 0.001) and season bins (R = 0.310, P = 0.001), while there was no significant difference between stations (R = 0.020, P = 0.075). The RPCA revealed a strong clustering of samples by depth (Figure 4A) across the PC1 axis (Figure 4C), whereas they are largely separated by season (Figure 4B) across the PC2 axis (Figure 4D). When plotting depth against PC1, one of the depth bins may need adjusting as there appears to be overlap between the communities for the 26-50 m and 51-72 m bins. In the seasonal bins, the winter samples generally cluster separately from the spring and summer samples. One notable exception is the deepest winter samples, which cluster with the other samples from the 400-500 m depth bin instead of the rest of the winter samples. When the bins are separated by month, we see high variation in community composition during September and March, which may indicate these are periods of transition.

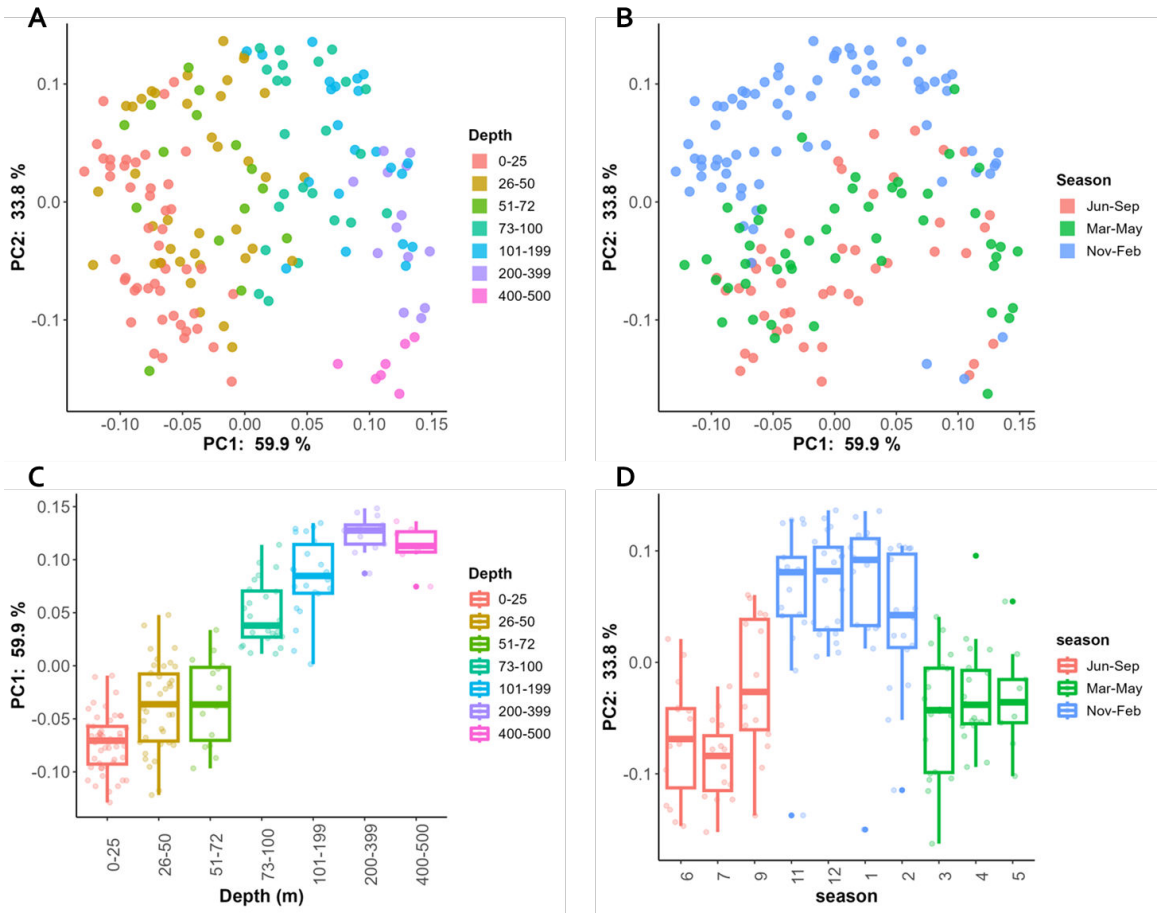

Figure 4. COI gene marker beta diversity results for all samples collected across all stations. The RPCA shows a relationship between (A) depth and PC1, while (B) PC2 is more related to seasonality. PC1 and PC2 are also plotted directly against (C) depth and (D) season, where the months are displayed chronologically from June 2021 to May 2022. All depths are reported in meters from the surface.

The top loading scores for PC1 include organisms in the class Hydrozoa and Phaeophyceae. Most of the larger scores were not categorized by MEGAN6, and when BLASTed manually they were most closely related to bacteria. The lowest loading scores driving PC1 are predominately diatoms in the class Mediophyceae, algae in the classes Haptophyta and Phaeophyceae, as well as two ASVs identified as the green algae species *Bathycoccus prasinos*. Plotting the relative abundance of each ASV versus depth revealed many of the ASVs with low loading scores, like the Haptophyte ASV_1, occupy the upper surface layers of the ocean (Figure 5A). In contrast, many of the high loading scores represent ASVs occupying the lower layers, like the Hydrozoa ASV_77 (Figure 5B).

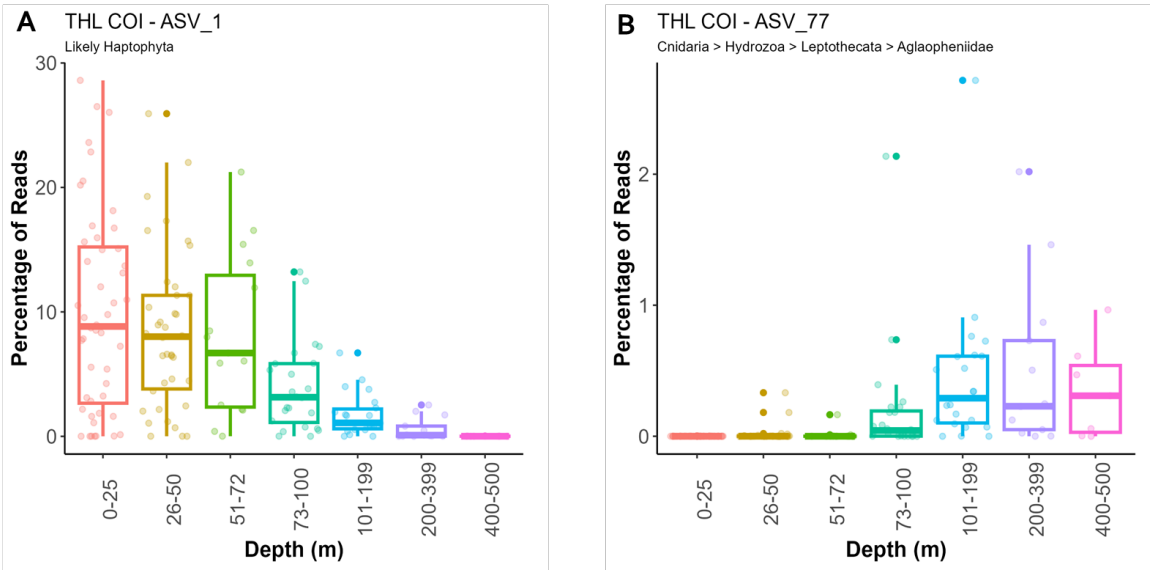

Figure 5. The percentage of total reads by depth (in meters) for (A) ASV_1 and (B) ASV_77. ASV_1 was not able to be assigned taxonomy by MEGAN6, however manually BLASTing the sequence against the NCBI GenBank database matched it most closely with organisms in the phylum haptophyta.

Looking at the loading scores for PC2, top scores are associated with algae in the classes Haptophyta, Phaeophyceae, Dictyochophyceae, and Oomycota. Although Oomycota share many characteristics with fungi, they are considered pseudofungi and more closely related to other algal groups like diatoms. The lowest loading scores are dominated by ASVs assigned to the copepod *Pseudocalanus mimus*, however other poorly identified arthropods are also well-represented, as well as Oomycota and diatoms in the class Mediophyceae. Plotting the relative abundance of each ASV by month reveals many of the ASVs with large loading scores are more abundant in the winter (Figure 6), while the lower loading scores are associated with an increase in abundance during the spring or summer months.

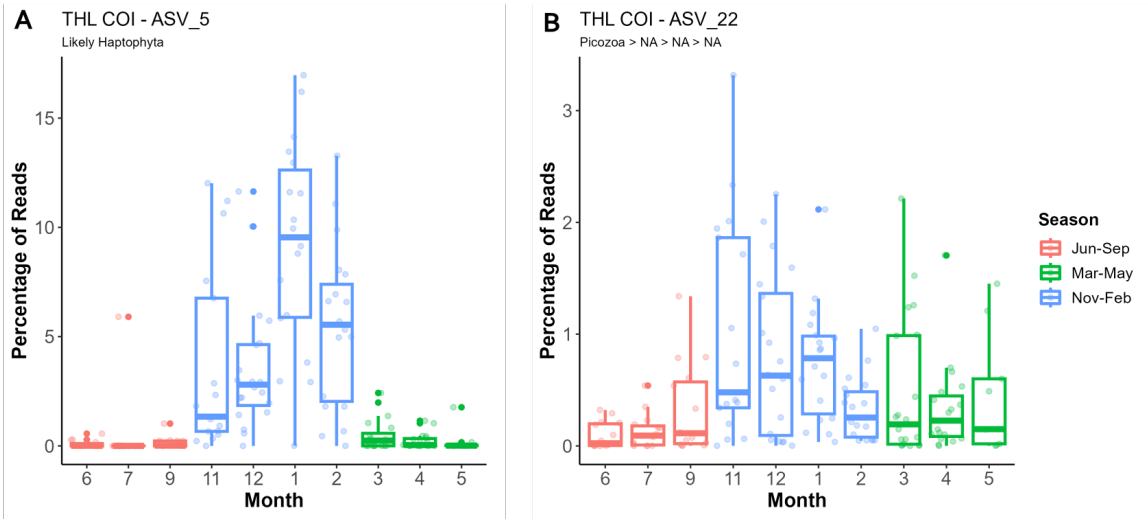

Figure 6. The percentage of total reads by month, which is further binned by season, for (A) ASV_5 and (B) ASV_22. Months are listed in chronological order from June 2021 to May 2022.

COMMUNITY DYNAMICS IN THE SURFACE LAYER FOR THE COI GENE

In the surface layer, the community composition was significantly different among season bins ($R = 0.667$, $P = 0.001$), while there was no significant difference among stations ($R = -0.041$, $P = 0.850$). This seasonal pattern is also strongly reflected among months ($R = 0.697$, $P = 0.001$), where the cold winter months are particularly distinct from the warm summer. The RPCA reveals a seasonal (and monthly) gradient along PC1, where the winter and summer months cluster at opposite ends of the ordination (Figure 7). PC2 may be driven by other seasonal environmental factors like temperature (Figure 7A) or upwelling index (Figure 7B), although there is still some uncertainty. However, when the temperature variable is separated into five equal bins and plotted against PC2, there appears to be a gradient from cold to warm, whereas for both upwelling indices the pattern is less clear.

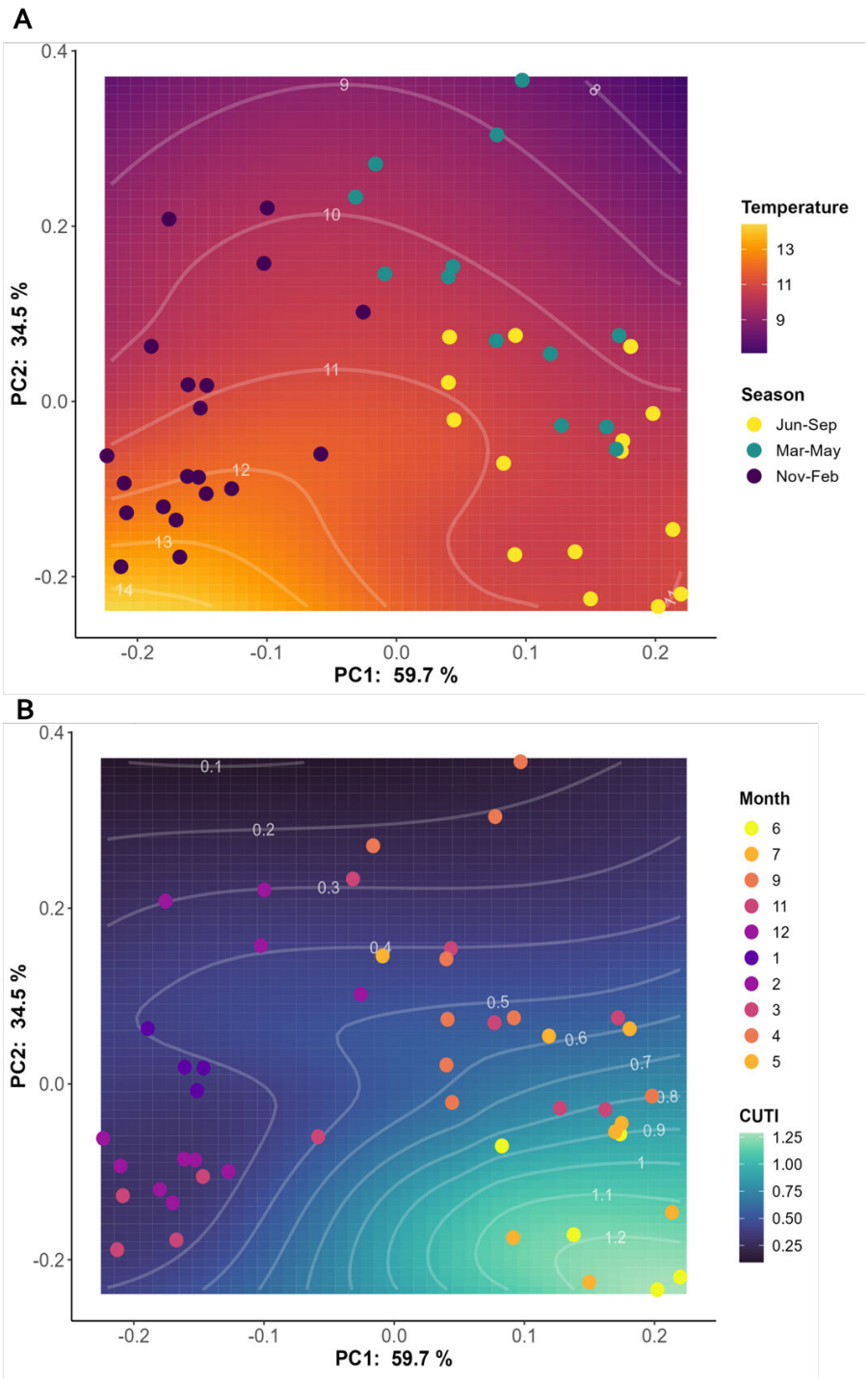

Figure 7. COI gene marker beta diversity results for samples collected from the surface layer (< 25 meters) across all stations. Both RPCA figures show a relationship between PC1 and season (or month), however PC2 is related to either (A) temperature or (B) CUTI index, which is displayed as a contour chart plotted behind the points. Temperature is reported in Celsius; CUTI is reported in m^2/s.

The seasonal differences in PC1 are driven predominately by ASVs assigned to the classes Haptophyta, Phaeophyceae, Mediophyceae, and Dictyochophyceae, which all have loading scores greater than 0.1. These classes are largely comprised heterokont protists in the phylum Gyrista that includes coccolithophores, brown algae, and diatoms. They were not assigned to Gyrista by the NCBI GenBank database, however the taxonomy appears to have been revised relatively recently in 2018 (Cavalier-Smith, 2018) and remains under debate. These ASVs were largely associated with increasing relative abundance during the winter months. ASVs with large negative values less than -0.1 were almost entirely comprised of the copepod *Pseudocalanus mimus*. When the relative abundances for all *P. mimus* ASVs are added together and plotted by month, it makes up a large percentage of reads predominately in the spring and summer months (Figure 8). Although the environmental factors driving PC2 are still unclear, diatoms, haptophytes, and chlorophytes are prevalent in loading scores higher than 0.1 and lower than -0.1. This includes the chlorophyte *Bathycoccus prasinos*, which is a photosynthetic picoplankton. Plotting these ASVs against environmental variables, as was done when all samples were combined, may help clarify the trends in PC2.

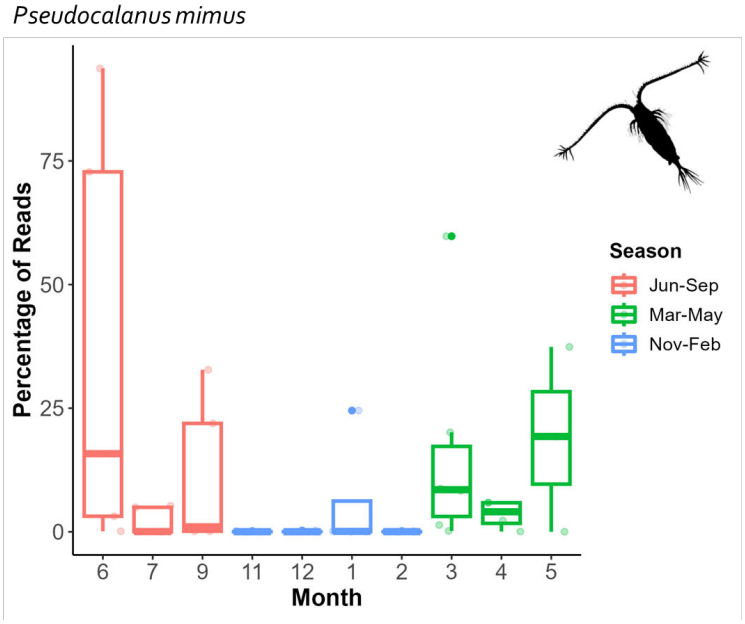

Figure 8. The sum of the percentage of reads for all ASVs assigned Pseudocalanus mimus, plotted by month and binned by season. Copepod silhouette provided by PhyloPic.

COMMUNITY DYNAMICS FOR THE 18S GENE

The overall patterns for the 18S gene are similar to what we see in the COI gene. For 18S, the community composition was significantly different between depth bins ($R = 0.419$, $P = 0.001$) and season bins ($R = 0.301$, $P = 0.001$). There was also a significant difference between stations ($R = 0.029$, $P = 0.045$), which was the most pronounced for stations TH01 and TH02, though there was still generally a lot of overlap in community composition between stations. The RPCA revealed a strong clustering of samples by depth (Figure 9A) across the PC1 axis (Figure 9C), whereas they are largely separated by season (Figure 9B) across the PC2 axis (Figure 9D). We used the same binning for the 18S gene as the COI gene, however each gene captures slightly different taxa and they may need to be adjusted. For example, the month of February does not cluster as neatly with the other winter months and the communities may be more representative of a spring transition.

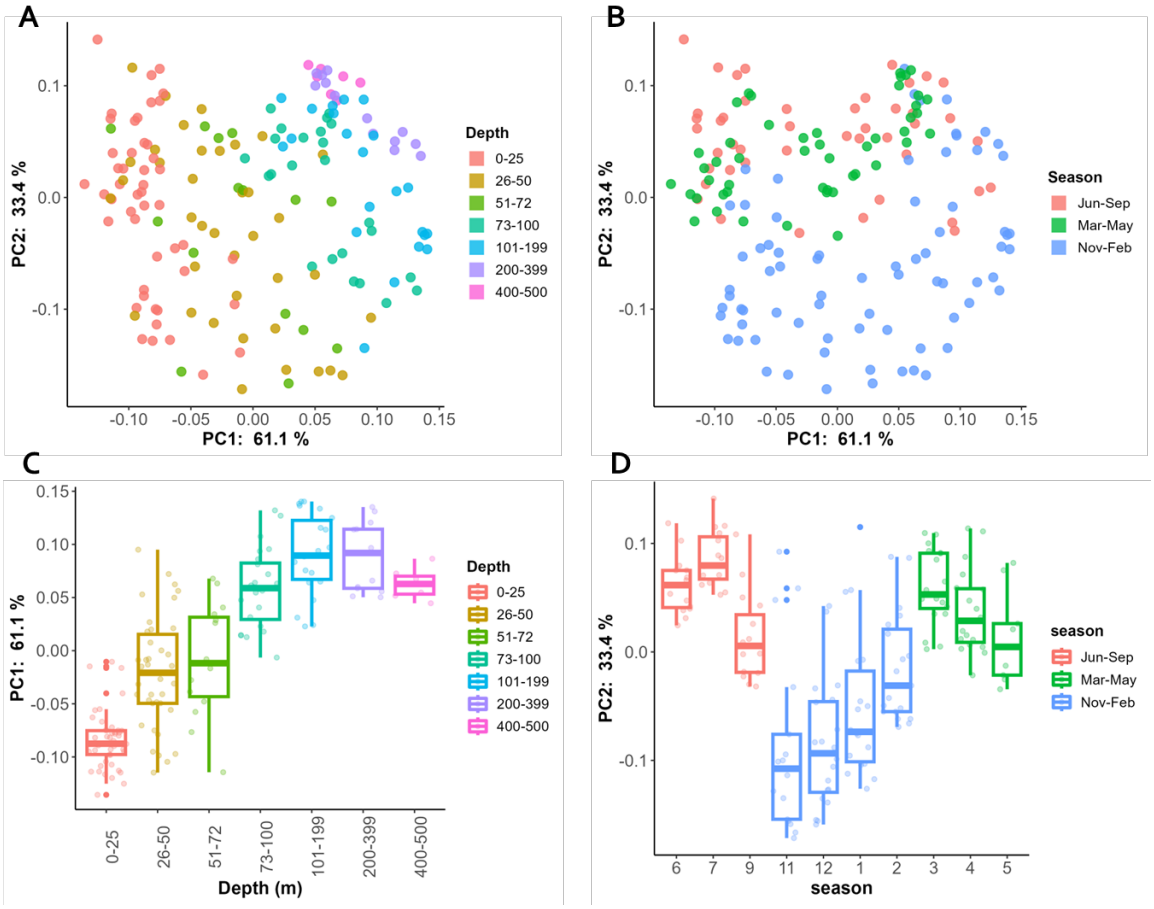

Figure 9. 18S gene marker beta diversity results for all samples collected across all stations. The RPCA shows a relationship between (A) depth and PC1, while (B) PC2 is more related to seasonality. PC1 and PC2 are also plotted directly against (C) depth and (D) season, where the months are displayed chronologically from June 2021 to May 2022. All depths are reported in meters from the surface.

Diatoms and dinoflagellates are of particular interest since they perform a large percentage of photosynthesis in the ocean (Nelson et al., 1995) and are often responsible for harmful algal blooms (*September 2021 Marine Biotoxin Monthly Report*, 2021). The diatom class Bacillariophyceae is most prevalent in the summer, with a strong peak occurring in September. In contrast, both Coscinodiscophyceae and Fragilarophyceae are in low abundance in the summer, increasing in late winter and peaking in the spring. Mediophycease is in low abundance through most of the year, with a large increase during the month of April. Alternatively, dinoflagellates (in the class Dinophyceae) are prevalent throughout most of the year, particularly in September, other than the spring when the drop substantially in April at the same time Coscinodiscophyceae and Mediophyceae are at their highest abundance. Overall, Dinophyceae is highly prevalent

in these samples, making up 15%-30% of all reads throughout most of the year (Figure 10).

Figure 10. Mean percentage of total reads over time for four classes of diatoms (Bacillariophyceae, Coscinodiscophyceae, Fragilariophyceae, and Mediophyceae) and one class of dinoflagellate (Dinophyceae). Percentages were averaged for each class within each month. Months are displayed in chronological order from June 2021 to June 2022.

Krill, which are important prey species for many larger organisms like fish and whales (Fisher et al., 2020), were selected for evaluating the similarities between the net tow and eDNA data. At the same time the eDNA water samples were collected, vertical net tows were performed across the upper 100 m of water for stations TH03, TH04, and TH05. The shape of the net tow data is similar across all three stations for the months that were sampled, where abundance is initially high, then dips dramatically in December, followed by an increase up through the spring months. For TH03 and TH05, the pattern is similar between the two methods, with decreases in abundance also primarily occurring during the winter. However, the pattern does not align in TH04, where there is a strong peak in March in the eDNA data, with very little krill detected in any other month. It is

notable that the highest monthly relative abundance at TH04 makes up 7.5% of the reads, which is much higher than the other stations where the percentage of reads are less than 0.1%.

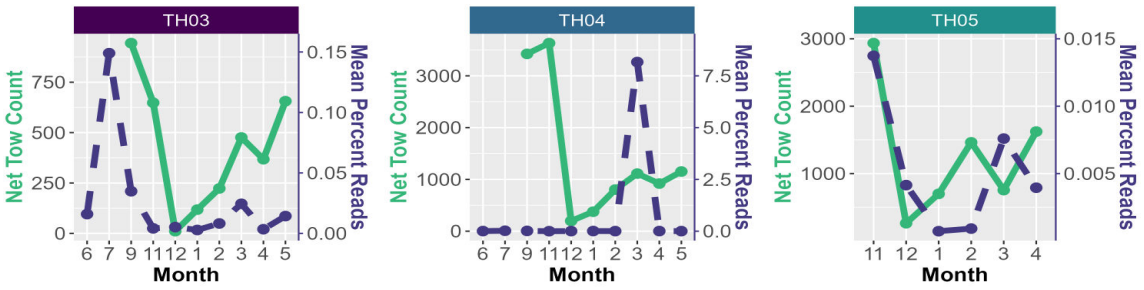

Figure 11. Seasonal patterns of net tow counts of krill (Family: Euphausiidae) in comparison to the mean percent reads of krill from eDNA. Only stations TH03 - TH05 are represented since they had the most complete data for both methods.

DISCUSSION

This study provides the first preliminary overview of pelagic community ecology in Northern California using environmental DNA. Community differences for both the COI and 18S genes were predominately driven by depth and season. Sampling efforts represent the upper 500 m of ocean, capturing a large gradient of light across depth. From the surface to 200 m of depth is the epipelagic zone where light is abundant, and many of the top loading scores are associated with photosynthetic organisms like diatoms, chlorophytes, and haptophytes. Haptophytes were not only in greater relative abundance towards the surface, but were also well-represented in the ASVs associated with seasonal differences, with many ASVs in higher relative abundance during the winter months. Since not all haptophyte ASVs are most abundant in the winter, these ASVs may represent a particular group or species that are able to take advantage of winter conditions. When BLASTed manually, ASV_5 matched most closely with *Emiliania huxleyi* and *Gephyrocapsa parvula*. *E. huxleyi* was onc of the dominant species in the autumn and winter communities on the west coast of the Iberian Peninsula, part of a large upwelling system. This site also sits at the mouth of a river that acts as a major source of nutrients during the rainy winter season (Vidal et al., 2017), while the THL lies just south of the Klamath River. Although

haptophytes make up a notable percentage of the plankton community in the California Current System, their seasonality is less well-studied than diatoms and dinoflagellates. This seasonal pattern has not been observed in Monterey Bay, where haptophytes are in the greatest abundance during April, May, and September (Bac et al., 2003).

Diatoms and dinoflagellates comprise the largest biomass of phytoplankton in the California Current System, however they display different seasonal patterns and typically bloom under different conditions (Fischer et al., 2020). Diatoms are responsible an estimated 35% of oligotrophic ocean primary productivity, up to 75% in nutrient-rich regions (Nelson et al., 1995), and they often dominate phytoplankton communities during upwelling events (Lassiter et al., 2006). In Bodega Bay (38°N), phytoplankton communities during May-June are dominated by a handful of genera in the classes Coscinodiscophyceae, Bacillariophyceae, and Fragilariophyceae, particularly the centric diatom *Chaetoceros* in the class Coscinodiscphyceae (Lassiter et al., 2006). Along the THL, Coscinodiscophyceae also displayed the greatest relative abundance in May-June out of all classes, however it is even more abundant during late winter before the upwelling season starts and peaked in April. It may be more informative to identify when peak chlorophyll concentrations are occurring at the surface and if it aligns with the abundance of diatoms, rather than upwelling indices alone, as well look at changes in genus-level abundance within Coscinodiscophyceae. Genera in Bacillariophyceae were also abundant during upwelling in Bodega Bay (Lassiter et al., 2006), and were the second most abundant diatom group which peaked towards the end of the upwelling season in September. Bacillariophyciae contains the genus *Pseudo-nitzchia*, which is a major causative agent in domoic acid toxicity and is surveyed for regularly along the California coast. Although most of the ASVs in our samples were assigned to the genus *Pseudo-nitzchia*, the California Department of Public Health reported low concentrations of *Pseudo-nitzchia* and domoic acid at Trinidad Pier (*September 2021 Marine Biotoxin Monthly Report*, 2021), though these samples are taken much closer to shore.

We found a high relative abundance of dinoflagellates throughout most of the year, including the month of July when the most water transport was occurring (though their populations are in the greatest abundance when turbulence is decreasing at the end of the upwelling season). While many dinoflagellate species contribute to primary productivity,

they are not typically considered a major part of the upwelling community in northern California. Summer surveys at both Bodega Bay (Lassiter et al., 2006) and Point Area (Chavez et al., 1991) found the abundance of dinoflagellates to be much lower that diatoms, however we find them in higher relative abundance using eDNA. This may be another reflection of an anomalous year, as large dinoflagellate blooms have been recorded in Monterey Bay when local conditions are particularly favorable (Fischer et al., 2020). Alternatively, it may be the result of dinoflagellates having large repetitive genomes, where a single individual may possess many copies of the 18S gene (Gong & Marchetti, 2019) that could artificially inflate relative abundance and make comparisons between taxa challenging. In spring 2022, the diatom-dinoflagellate pattern was more typical, with the highest relative abundance of diatoms in April coinciding with a low abundance of dinoflagellates.

Copepods and krill represent two major groups of zooplankton that are ecologically significant in this region. The presence of different copepod species can be used as a seasonal indicator in the California Current System as different species show strong habitat preferences, particularly for water temperature (Bjorkstedt & Peterson, 2015). In the surface community, the copepod *Pseudocalanus mimus* made up a large percentage of reads during spring and summer, particularly at one station in June where it represented over 75% of all reads. *P. mimus* is a large lipid-rich species with an affinity for cold-water that typically appears in the northern portion of the California Current System concurrently with the start of the upwelling season (Bjorkstedt & Peterson, 2015), which aligned with our results. The seasonal pattern of krill (i.e. Euphausiids) did not quite meet expectations for each station. The net tow and eDNA data for station TH03 (which occurs just past the shelf break) and the farthest station TH05 were generally in alignment, with the lowest abundances occurring in December. For TH04, the seasonal pattern of net tow data was similar to the other stations, however the eDNA data showed a huge peak only in March. The relative abundance is so high in comparison to the rest of the krill eDNA results that it should be looked at more closely to determine if an error has occurred. Like copepods, Euphausiids display species-specific preferences for different environmental conditions (Marinovic et al., 2002), so it may also be beneficial to re-analyze seasonal patterns at the

species level (using the COI gene instead of 18S, which is better at assigning species taxonomy for krill).

This dataset only covers a one-year period where upwelling was particularly strong during the spring and summer, which may not be representative of the average seasonal pattern for this region. June 2021 through May 2022 occurred during a La Niña phase, where anomalous upwelling was predicted to occur in conjunction with lower temperatures, higher salinities, and an abundance of chlorophyll. While most of the California Current System experienced upwelling only slightly above average, the middle latitudes (between 36°N and 42°N) exhibited highly anomalous water transport throughout most of the year, coupled with slightly cooler sea surface temperatures, high salinities, and high concentrations of chlorophyll. Community composition is expected to shift during La Niña conditions, not only with phytoplankton blooms but also increasing abundances of northern copepods species and krill, which provide a rich food source for young of the year fish (Thompson et al., 2022). The addition of subsequent years of eDNA data for the THL line will likely reveal interesting interannual variation, as well as fill in the gaps for August and October so that the monthly transitions can be better defined.

Although metabarcoding methods have become more robust in recent years, there are still major areas for improvement that affect our results. Taxonomic assignment relies on having a comprehensive database of identified barcodes, however many marine organisms are missing barcodes and a high percentage of ASVs could not be assigned even at the phylum level. This has a bigger impact on our interpretation of the results when those ASVs are associated with large loading scores in the PCA analysis. For example, eight out of ten of the top largest loading scores driving PC1 for the COI gene could not be assigned using MEGAN6. When BLASTed manually against the NCBI database, five of those ASVs were revealed to most likely be bacteria, which may be affecting the differences between depths and should have been removed from the final analysis. There was also an issue with ASVs in the phylum Arthropoda being assigned to terrestrial organisms like spiders and flies. Although we might expect some DNA from terrestrial organisms to appear in our sequences, the large number of ASVs across different samples suggests this was more likely resulting from misclassification. Misclassification is a common issue for lowest common ancestor algorithms using global databases because it assumes all

sequences are valid, even if they are not relevant (like terrestrial spiders) or are from a different geographic areas (Curd et al., 2019). Adjusting the parameters guiding the MEGAN6 algorithm may help resolve these issues. In addition, a curated database could be constructed that is specific to both this environment and geographic region, which may provide more accurate taxonomic assignments than a database that is global and all-inclusive (Gold et al., 2021).

This overview of the Trinidad Head Line eDNA dataset provides a solid starting point for further analysis. Although the 12S gene was sequenced for this time series, an analysis of the vertebrate community has not yet been performed. In addition, while the surface community was analyzed for the COI gene, it has not yet been analyzed for the 18S, which is important since 18S typically detects a broader diversity of photosynthetic phytoplankton (Sawaya et al., 2019) that are likely to inhabit the surface layer. Though we did not observe strong differences between nearshore and offshore stations as we might have expected (especially given the differences in depths sampled), this may be more pronounced when looking at individual seasons, particularly summer when upwelling should stimulate more primary productivity at stations closer to the coastline. Additionally, we can compare the seasonal shifts in major taxa of interest like diatoms, dinoflagellates, and copepods to the eDNA time series conducted monthly in Monterey Bay. Although both regions experience strong seasonal upwelling, they contain very different geographic features and ongoing monitoring will help reveal if this is an important component to community composition. Finally, the Humboldt region has been identified as a potential site for offshore windfarms (Raghukumar et al., 2023), making the analysis of additional years of data and continued biomonitoring even more critical for accurately predicting potential ecosystem impacts.

CONCLUSIONS/RECOMMENDATIONS

The Trinidad Head Line lies in an important transition zone between the northern and southern portions of the California Current System, with unique geographic features like Cape Blanco and Cape Mendocino that force cool upwelled waters far offshore. Ongoing monitoring of this region will help detect populations shifting up the coast in

response to warming, as well as identify ecological factors like nutrient concentrations that are expected to impact the productivity of local fisheries. These efforts are necessary to continually update management policies so they are flexible enough to respond to our changing oceans, ensuring there productive ecosystems remain as healthy as possible for years to come.

ACKNOWLEDGEMENTS (Normal, Times New Roman, 12 pt, bold)

The MBARI internship is made possible through the Dean and Helen Witter Family Fund, the Rentschler Family Fund, the David and Lucile Packard Foundation, and the Maxwell/Hanrahan Foundation. I owe an immense amount of gratitude to Jacoby Baker for his mentorship throughout this project, as well as Dr. Francisco Chavez and Dr. Kathleen Pitz for their expertise in working with oceanographic and eDNA data. I would also like to thank Dr. Eric Bjorkstedt's lab group for their ongoing efforts in monitoring the Trinidad Head Line, which included collecting and processing the eDNA samples, as well as generating the net tow data for comparing the abundances between methods.

References:

References (Normal, Times New Roman, 12 pt). Format should match the format that can be found online at http://www.mbari.org/products/publications/

Bac, M. G., Buck, K. R., Chavez, F. P., & Brassell, S. C. (2003). Seasonal variation in alkenones, bulk suspended POM, plankton and temperature in Monterey Bay, California: Implications for carbon cycling and climate assessment. *Organic Geochemistry*, *34*(6), 837–855. https://doi.org/10.1016/S0146-6380(02)00248-6

Bailey, J., Durbin, E. G., & Rynearson, T. (2016). Species composition and abundance of copepods in the morphologically cryptic genus Pseudocalanus in the Bering Sea. *Deep Sea Research Part II: Topical Studies in Oceanography*, *134*, 173–180. https://doi.org/10.1016/j.dsr2.2015.04.017

Barth, J. A., Pierce, S. D., & Smith, R. L. (2000). A separating coastal upwelling jet at Cape Blanco, Oregon and its connection to the California Current System. *Deep Sea Research Part II: Topical Studies in Oceanography*, *47*(5), 783–810. https://doi.org/10.1016/S0967-0645(99)00127-7

Bjorkstedt, E. P., & Peterson, W. T. (2015). Zooplankton Data from High-Frequency Coastal Transects. In *Coastal Ocean Observing Systems* (pp. 119–142). Elsevier. https://doi.org/10.1016/B978-0-12-802022-7.00008-0

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, *37*(8), 852–857. https://doi.org/10.1038/s41587-019-0209-9

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Cavalier-Smith, T. (2018). Kingdom Chromista and its eight phyla: A new synthesis emphasising periplastid protein targeting, cytoskeletal and periplastid evolution, and ancient divergences. *Protoplasma*, *255*(1), 297–357. https://doi.org/10.1007/s00709-017-1147-3

Chavez, F. P., Barber, R. T., Kosro, P. M., Huyer, A., Ramp, S. R., Stanton, T. P., & Rojas de Mendiola, B. (1991). Horizontal transport and the distribution of

nutrients in the Coastal Transition Zone off northern California: Effects on primary production, phytoplankton biomass and species composition. *Journal of Geophysical Research: Oceans*, *96*(C8), 14833–14848. https://doi.org/10.1029/91JC01163

Chavez, F. P., & Messié, M. (2009). A comparison of Eastern Boundary Upwelling Ecosystems. *Progress in Oceanography*, *83*(1), 80–96. https://doi.org/10.1016/j.pocean.2009.07.032

Cheng, L., von Schuckmann, K., Abraham, J. P., Trenberth, K. E., Mann, M. E., Zanna, L., England, M. H., Zika, J. D., Fasullo, J. T., Yu, Y., Pan, Y., Zhu, J., Newsom, E. R., Bronselaer, B., & Lin, X. (2022). Past and future ocean warming. *Nature Reviews Earth & Environment*, *3*(11), Article 11. https://doi.org/10.1038/s43017-022-00345-1

Curd, E. E., Gold, Z., Kandlikar, G. S., Gomer, J., Ogden, M., O'Connell, T., Pipes, L., Schweizer, T. M., Rabichow, L., Lin, M., Shi, B., Barber, P. H., Kraft, N., Wayne, R., & Meyer, R. S. (2019). Anacapa Toolkit: An environmental DNA toolkit for processing multilocus metabarcode datasets. *Methods in Ecology and Evolution*, *10*(9), 1469–1475. https://doi.org/10.1111/2041-210X.13214

Didion, J. P., Martin, M., & Collins, F. S. (2017). Atropos: Specific, sensitive, and speedy trimming of sequencing reads. *PeerJ*, *5*, e3720. https://doi.org/10.7717/peerj.3720

Fischer, A. D., Hayashi, K., McGaraghan, A., & Kudela, R. M. (2020). Return of the "age of dinoflagellates" in Monterey Bay: Drivers of dinoflagellate dominance examined using automated imaging flow cytometry and long-term time series

analysis. *Limnology and Oceanography*, *65*(9), 2125–2141. https://doi.org/10.1002/lno.11443

Fisher, J. L., Menkel, J., Copeman, L., Shaw, C. T., Feinberg, L. R., & Peterson, W. T. (2020). Comparison of condition metrics and lipid content between Euphausia pacifica and Thysanoessa spinifera in the northern California Current, USA. *Progress in Oceanography*, *188*, 102417. https://doi.org/10.1016/j.pocean.2020.102417

Gold, Z., Curd, E. E., Goodwin, K. D., Choi, E. S., Frable, B. W., Thompson, A. R., Walker Jr., H. J., Burton, R. S., Kacev, D., Martz, L. D., & Barber, P. H. (2021). Improving metabarcoding taxonomic assignment: A case study of fishes in a large marine ecosystem. *Molecular Ecology Resources*, *21*(7), 2546–2564. https://doi.org/10.1111/1755-0998.13450

Gong, W., & Marchetti, A. (2019). Estimation of 18S Gene Copy Number in Marine Eukaryotic Plankton Using a Next-Generation Sequencing Approach. *Frontiers in Marine Science*, *6*. https://www.frontiersin.org/articles/10.3389/fmars.2019.00219

Harvey, J. B. J., Johnson, S. B., Fisher, J. L., Peterson, W. T., & Vrijenhoek, R. C. (2017). Comparison of morphological and next generation DNA sequencing methods for assessing zooplankton assemblages. *Journal of Experimental Marine Biology and Ecology*, *487*, 113–126. https://doi.org/10.1016/j.jembe.2016.12.002

Hickey, B. M., & Banas, N. S. (2008). Why is the Northern End of the California Current System So Productive? *Oceanography*, *21*(4), 90–107.

Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.-J., & Tappu, R. (2016). MEGAN Community Edition—Interactive Exploration

and Analysis of Large-Scale Microbiome Sequencing Data. *PLOS Computational Biology*, *12*(6), e1004957. https://doi.org/10.1371/journal.pcbi.1004957

Jacox, M. G., Edwards, C. A., Hazen, E. L., & Bograd, S. J. (2018). Coastal Upwelling Revisited: Ekman, Bakun, and Improved Upwelling Indices for the U.S. West Coast. *Journal of Geophysical Research: Oceans*, *123*(10), 7332–7350. https://doi.org/10.1029/2018JC014187

Lamy, T., Pitz, K. J., Chavez, F. P., Yorke, C. E., & Miller, R. J. (2021). Environmental DNA reveals the fine-grained and hierarchical spatial structure of kelp forest fish communities. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-93859-5

Lassiter, A. M., Wilkerson, F. P., Dugdale, R. C., & Hogue, V. E. (2006). Phytoplankton assemblages in the CoOP-WEST coastal upwelling area. *Deep Sea Research Part II: Topical Studies in Oceanography*, *53*(25), 3063–3077. https://doi.org/10.1016/j.dsr2.2006.07.013

Mackas, D. L., Strub, P., Thomas, A., & Montecino, V. (2006). Eastern ocean boundaries pan-regional overview. *The Sea*, 21–59.

Marinovic, B. B., Croll, D. A., Gong, N., Benson, S. R., & Chavez, F. P. (2002). Effects of the 1997–1999 El Niño and La Niña events on zooplankton abundance and euphausiid community composition within the Monterey Bay coastal upwelling system. *Progress in Oceanography*, *54*(1), 265–277. https://doi.org/10.1016/S0079-6611(02)00053-8

Martino, C., Morton, J. T., Marotz, C. A., Thompson, L. R., Tripathi, A., Knight, R., & Zengler, K. (2019). A Novel Sparse Compositional Technique Reveals Microbial

Perturbations. *mSystems*, *4*(1), 10.1128/msystems.00016-19. https://doi.org/10.1128/msystems.00016-19

Min, M. A., Needham, D. M., Sudek, S., Truelove, N. K., Pitz, K. J., Chavez, G. M., Poirier, C., Gardeler, B., von der Esch, E., Ludwig, A., Riebesell, U., Worden, A. Z., & Chavez, F. P. (2023). Ecological divergence of a mesocosm in an eastern boundary upwelling system assessed with multi-marker environmental DNA metabarcoding. *Biogeosciences*, *20*(7), 1277–1298. https://doi.org/10.5194/bg-20-1277-2023

Muller-Karger, F. E., Kavanaugh, M. T., Montes, E., Balch, W. M., Breitbart, M., Chavez, F. P., Doney, S. C., Johns, E. M., Letelier, R. M., Lomas, M. W., Sosik, H. M., & White, A. E. (2014). A Framework for a Marine Biodiversity Observing Network Within Changing Continental Shelf Seascapes. *Oceanography*, *27*(2), 18–23. https://doi.org/10.5670/oceanog.2014.56

NCBI Resource Coordinators. (2018). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, *46*(D1), D8–D13. https://doi.org/10.1093/nar/gkx1095

Nelson, D. M., Tréguer, P., Brzezinski, M. A., Leynaert, A., & Quéguiner, B. (1995). Production and dissolution of biogenic silica in the ocean: Revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Global Biogeochemical Cycles*, *9*(3), 359–372. https://doi.org/10.1029/95GB01070

Raghukumar, K., Nelson, T., Jacox, M., Chartrand, C., Fiechter, J., Chang, G., Cheung, L., & Roberts, J. (2023). Projected cross-shore changes in upwelling induced by

offshore wind farm development along the California coast. *Communications Earth & Environment*, *4*(1), Article 1. https://doi.org/10.1038/s43247-023-00780-y

Robertson, R. R., & Bjorkstedt, E. P. (2020). Climate-driven variability in Euphausia pacifica size distributions off northern California. *Progress in Oceanography*, *188*, 102412. https://doi.org/10.1016/j.pocean.2020.102412

Sawaya, N. A., Djurhuus, A., Closek, C. J., Hepner, M., Olesin, E., Visser, L., Kelble, C., Hubbard, K., & Breitbart, M. (2019). Assessing eukaryotic biodiversity in the Florida Keys National Marine Sanctuary through environmental DNA metabarcoding. *Ecology and Evolution*, *9*(3), 1029–1040. https://doi.org/10.1002/ece3.4742

September 2021 Marine Biotoxin Monthly Report (Technical Report 21–21). (2021). California Department of Public Health. https://storymaps.arcgis.com/stories/7ffc098badde49eda318406ce5609b6e

Thompson, A. R., Bjorkstedt, E. P., Bograd, S. J., Fisher, J. L., Hazen, E. L., Leising, A., Santora, J. A., Satterthwaite, E. V., Sydeman, W. J., Alksne, M., Auth, T. D., Baumann-Pickering, S., Bowlin, N. M., Burke, B. J., Daly, E. A., Dewar, H., Field, J. C., Garfield, N. T., Giddings, A., … Weber, E. D. (2022). State of the California Current Ecosystem in 2021: Winter is coming? *Frontiers in Marine Science*, *9*. https://www.frontiersin.org/articles/10.3389/fmars.2022.958727

Vidal, T., Calado, A. J., Moita, M. T., & Cunha, M. R. (2017). Phytoplankton dynamics in relation to seasonal variability and upwelling and relaxation patterns at the

mouth of Ria de Aveiro (West Iberian Margin) over a four-year period. *PLOS ONE*, *12*(5), e0177237. https://doi.org/10.1371/journal.pone.0177237