Variation in Zooplankton Community Composition in Prince William Sound across Space and Time

Final Report to Prince William Sound Regional Citizens' Advisory Council (PWSRCAC)

Coastal Disease Ecology Laboratory
Smithsonian Environmental Research Center

Dr. Katrina M Pagenkopp Lohan
Ruth DiMaria

&

Molecular Ecology Laboratory
Moss Landing Marine Laboratory

Dr. Jonathan Geller

Report Revised: July 5, 2022
Report Submitted: May 16, 2022

EXECUTIVE SUMMARY

Creating comprehensive species lists for benthic marine habitats typically require costly and laborious large-scale collections of samples, exhaustive sorting of specimens, and expert taxonomic identification. When time, labor, expertise, or funds are limiting, an alternative approach can be collection and genetic analysis of planktonic larvae of bottom-dwelling species (referred to as meroplankton) in the water column. This approach may also be well-suited to detect nonindigenous species (NIS), as many of these become established after transport in ballast water as larval stages. Metabarcoding is the simultaneous genetic analysis of the same gene from individual organisms from multiple species in an environmental sample of biomass, in this case from plankton samples. In metabarcoding, individual DNA sequences are grouped by similarity into clusters called Operational Taxonomic Units (OTU) that represent biological species, which can be assigned taxonomic names through comparisons to sequences in established databases. In previous studies, we employed this approach to describe zooplankton communities in Port Valdez, but we lacked information on the variability of zooplankton communities that is necessary to optimize a sampling program. In the current study, we used DNA metabarcoding to examine the potential sources of variation (namely season, tide, daylight, and sampling location) for zooplankton community richness (defined as the number of species in a community) and composition (defined as the proportion of each species in the community) in the Port Valdez. In doing so, we hoped to inform improved sampling strategies and better understand prior results. In this study, our results showed high OTU diversity, with sequences from a few species dominating the samples. A spring to summer shift in the zooplankton

community was observed, consistent with known zooplankton dynamics in Port Valdez. Variation in community composition was primarily attributed to date of sampling and not to location, day/night cycles or tidal stage. Finally, some taxa expected from fouling communities (defined as the biotic growth on hard surfaces), which are often rich with NIS, were underrepresented in these results. Additionally, some species found by morphological identification of specimens (using physical characteristics such as shape, size and color) in fouling communities in an earlier study did not appear in our results. Conversely, many of the taxa found in this survey were not reported in the morphological survey. In retrospect, fouling communities are a small fraction of the total benthic habitat in Port Valdez, and their larvae may similarly be a small component of the total zooplankton community. Based on these results, we make the following recommendations for future surveys: 1) consider increased sequencing depth or molecular strategies to suppress dominant species to enhance detection of fouling species, 2) increase replication of summer sampling to increase potential detection of meroplankton, and 3) utilize a hybrid strategy to directly sample fouling communities, such as conducting complementary, simultaneous morphological and metabarcode surveys.

INTRODUCTION

Sampling of plankton communities is a novel approach to monitoring benthic (defined as bottom-dwelling) marine communities when planktonic larvae of benthic species (referred to as meroplankton) are present in the water column. Diversity in plankton samples is also easier to measure compared to sampling the benthic communities. Metabarcoding is defined as the exhaustive sequencing of species-diagnostic genomic fragments from DNA extractions of bulk community samples. Metabarcoding of plankton is well-suited to detect nonindigenous species (NIS) that became established after transport in ballast water because these are biased toward species with planktonic larvae (Carlton & Geller 1993). From 2016 to 2019, we applied a metabarcoding approach to plankton communities in Port Valdez, Alaska, with the primary goal of detecting NIS; however, few NIS were seen in those datasets. While detection of planktonic larvae depends on prior adult reproduction, which is driven by adult environmental physiology, other factors such as local circulation, tidal patterns, and larval behaviors can also impact planktonic larvae richness and spatiotemporal variation. Thus, we were concerned that our prior studies under-sampled plankton in Port Valdez, as the limited sampling from a single date in a few locations may have failed to collect many species actually present in the benthic communities.

The current study was undertaken to examine potential sources of variation (specifically season, tide, daylight, and location) for the estimation of zooplankton community richness and composition in the Port Valdez. In doing so, we hoped to inform improved sampling strategies and to better understand results from our prior studies. We proposed a sampling design that would spread effort among days, weeks, and months to assess variation at these time scales. We included samples from three nearshore areas in Port Valdez to assess spatial variation. We also included daytime and nighttime sampling on some days in one site because plankton are known to exhibit phototaxis (i.e., bodily movement in response to light, either toward or away from the source). Finally, we sampled at different times in the tidal cycle in one site that was near the drainage of an extensive mudflat to explore potential habitat related differentiation in plankton composition. Several sampling schemes were considered, and the implemented plan reflected limitations of staffing, accessibility, and cost (Table 1).

METHODS

Sample collection

Zooplankton samples were collected from Prince William Sound, Alaska, from April through September in 2021 from three locations: Valdez Harbor (VDZ), the Container Terminal (CON), and the Valdez Marine Terminal (VMT) (Table 1). Tow samples were collected using a weighted plankton net (80 µm mesh, 0.5 m diameter) deployed to 5 m depth (except where the depth was less than 5 m in which case the net was lowered but not allowed to stir the bottom) and pulled vertically up through the water column. Three replicates were collected at each location per sampling event, assigned a unique ID, preserved, and shipped to the Coastal Disease Ecology Laboratory in Edgewater, Maryland, for metabarcoding and analyses.

DNA extraction, PCR amplification, and sequencing

Genomic DNA was extracted from a subset of the zooplankton collected from each replicate. Negative extraction controls were included to identify potential contaminants in the library preparation. A portion of the mitochondrial COI gene was amplified using primers fbLCOF1 (J.

Geller, unpublished) and jgHCO2198 (Geller *et al.* 2013). This COI gene fragment is a genetic marker, or "DNA barcode," commonly used to identify animals and so is well represented in public databases to aid taxonomic assignment of DNA sequences. All Polymerase Chain Reaction (PCR) reactions were generated in triplicate to mitigate potential variation across replicates in PCR. Specific DNA tags were added to the beginning and end of the PCR products as indices to later identify the source sample for each DNA sequence. The sequences were then purified to remove small and spurious fragments. The concentration of DNA per sample was then quantified. Based on those calculations, DNA from each sample was then pooled based on equimolar concentrations into three libraries for sequencing, with the intent of having the same concentration of DNA lead to a similar number of sequences per sample. The final pooled libraries were sequenced using a MiSeq v3 600 Reagent Kit (Illumina) on an Illumina MiSeq platform at the Laboratories of Analytical Biology at the Smithsonian National Museum of Natural History.

Bioinformatics

Sequences from all three runs were combined for bioinformatic analyses. Primer sequences were removed. Sequences were quality trimmed, merged, and chimeras (an artifact where partial PCR products from different species can be joined) were removed using the DADA2 package (Callahan *et al.* 2016) in R (Team 2020). Summary statistics were generated using the phyloseq (McMurdie & Holmes 2013) and vegan (Okasanen *et al.* 2014) packages in R. Individual sequences (also referred to as reads) were clustered at a 95% similarity threshold to form Operational Taxonomic Units (OTU), which were the unit used for community analysis. OTUs are treated as a proxy for biological species. To specifically look at temporal changes at each location, compare across locations, compare day vs. night, and compare across the tidal cycle, samples were parsed into different datasets to ensure an even sample size for all comparisons (Table 2). To assign taxonomic names to OTUs, a representative sequence from each OTU was compared first to a private MLML COI Database and then to the publicly available GenBank nucleotide (nr) database using BLAST (Altschul 1990). We annotated those OTUs that had an e-value of $\leq 1x10^{-30}$, $\geq$95% pairwise identify, and $\geq$90% pairwise coverage (or overlap) to a database record. If discrepancies existed, then the identification from the MLML database was given priority. The worms package (Chamberlain 2018) in R was used to add uniform upstream taxonomy for those taxa with matches in the World Register of Marine Species (WoRMs) database. Graphs were created to show species richness, taxonomic composition, and community similarity across different factors. Additionally, PERMANOVAs were conducted to see which factors were statistically significant in differentiating zooplankton communities.

The global geographic distributions of all OTUs that could be assigned a binomial name (genus and species) were mapped using records in the OBIS database (OBIS 2002). OBIS is a database of species distributions based on physical collections associated with museums and universities. As such, it does not include records based solely on appearance in the literature. Too, not all physical collections have sent data to OBIS. As in any species database, taxonomic accuracy in OBIS is likely imperfect, which could distort the reported distribution of some species. Therefore, OBIS should not be considered definitive of species distributions. Bearing in mind these caveats, maps were examined by eye to suggest potential NIS, which were those species with disjunct distributions that do not conform to provincial concepts of biogeography. Species tagged as potential NIS in Alaska should be referred to taxonomic experts for further evaluation.

RESULTS

In total, 47,540,396 raw reads were generated, which was reduced to 31,208,592 reads after initial filtering, merging, and chimera removal. With the removal of negative control samples, 31,206,244 reads remained for comparative analyses with 1,257 OTUs (approximations for species-level comparisons across sequence data) (Table 2). When all OTUs that could not be identified to the Kingdom Animalia by BLAST were removed from the dataset, 78% of the reads (n = 24,447,209) were assigned to animals, resulting in 195 OTUs (Total_Animal dataset; Table 2). After parsing the different datasets for statistical comparisons, all datasets contained over 1 million sequences, with the VMT dataset having the least number of samples, the least number of sequences, and the least number of OTUs, as expected (Table 2).

I. Comparisons across sampling locations

Alpha diversity (species richness)

For examining species richness (defined as the number of different species present in a particular sample), when we were not statistically comparing across a factor, all samples collected at all sites were included. When statistical comparisons were being made to tease out factors driving zooplankton richness or community composition, then the All3 dataset (Table 2), containing equal numbers of samples collected from the same months from all three sites, was used.

For this analysis, OTUs were generated to approximate species. Alpha diversity metrics using OTU richness were assessed using the Chao1 diversity metric, which is a nonparametric method that incorporates abundance into richness estimates as rare OTUs are presumed the most important in assessing how many additional taxa are missing. Our results indicated that alpha diversity varied across locations and months sampled (Figure 1). When examining all the data from all samples (parsing the Total_Animals dataset by location; Figure 1), alpha diversity was highest at the Container Dock and the Valdez Harbor in May, but highest at the Valdez Marine Terminal in April. When comparing the alpha diversity metrics for the All3 dataset (Figure 2), at the Container Dock, OTU richness was highest in July, then similar across other months. In contrast, at Valdez Harbor and the Valdez Marine Terminal, OTU richness was highest in April and lowest in May and August at the Valdez Harbor, but lowest in May and June at the Valdez Marine Terminal.

OTU accumulation curves were created to examine if the sampling effort both overall and across sites appeared sufficient for capturing all species likely present at those sites. For these analyses, all the animal OTUs across months were combined, for a broad view at the number of species at each site across the sampling time frame (Total_Animals dataset). If the sampling effort was sufficient to capture all the OTU richness at a site, then these curves would eventually flatten out to straight lines (in other words, they would reach an asymptote), indicating that adding more samples would not result in the addition of new taxa to the dataset. Across the four datasets examined, the accumulation curves do not appear to reach an asymptote at any of the three sites sampled (Figure 3 B-D), nor do all the samples combined appear to asymptote (Figure 3A). This indicates that OTU richness across these sites is high and additional sampling would be required to capture the total animal richness at these sites from April to September.

Beta diversity

To examine the similarity in community composition (defined as the contribution of each species to the total community) temporally within a site and across the sites, we created multidimensional scaling plots, either a non-metric multidimensional scaling (NMDS) plot or a Principal Coordinates Analysis (PCoA) plot. Both types of plots take a distance matrix as input, then condense the multiple factors present into a 2-dimensional space. In a PCoA, multiple eigenvalues and eigenvectors are calculated, ranked from greatest to highest, and the top two are used to plot the data into 2-dimensional space. In a NMDS, the method is non-metric, as it converts the dissimilarity values into ranks, which are then used for the iterative calculation performed. In both types of plots, the closer two points are to each other, the more similar they are. Thus, points that are closer together in these graphs indicate that the community composition in those samples is similar. The two axes plotted for the PCoA (Figure 4) account for 33.7% of the variation across the Total_Animals dataset. The PCoA plots generated by parsing the Total_Animals dataset by location indicate that the community composition in samples collected from all three sites in April and May are both different from each other and different from the communities collected during other months (Figure 4). At all three sites, samples collected from June through September cluster closely together and the ellipses overlap, indicating that the composition of these samples is highly similar.

To further explore how the timing of sampling impacts the community composition, we created the All3 dataset, containing the same number of samples across months across sites. The NMDS plot with this dataset (Figure 5) shows that samples collected in April across all sites are more similar in composition to each other and distinct from the community composition in samples collected from all three sites during the subsequent months. Additionally, there appears to be little differentiation in community composition at any of the three sites from May to September, indicating that these communities are similar across this time frame, regardless of from where the samples were taken.

We then conducted a PERMANOVA to compare the community composition in the All3 dataset to see if month or location were statistically significant factors. The PERMANOVA compares groups of objects (in this case groups of metazoan zooplankton) to test the null hypothesis that the centroid location and dispersion of those groups are equivalent for all groups. A rejection of the null hypothesis indicates that either the location of the centroid and/or the spread of the objects (also referred to as the dispersion) is different between the groups. We then conducted a post hoc test, the Tukey test, to determine if the spread of the objects is significantly different. When this test is significant it indicates that there is a dispersion event, and there may or may not also be actual differences in the centroids between groups. In this case, the PERMANOVA results for the All3 dataset indicated that location did not have a significant impact on community composition ($p = 0.122$, All3 – Location; Table 3), but month did ($p = 0.001$, All3-Month; Table 3). The Tukey test for the All3-Month, indicated that there is a dispersion event ($p = <0.0001$; All3-Month; Table 4), which is evident given the spread of samples in the NMDS plots. Combining the output from the NMDS plot and these results, it appears that communities shifted across months with different degrees of dispersion.

Taxa
The taxa identified included animals from eight phyla (Figure 6, Appendix A). By far the most abundant, based on the number of sequences, were the arthropods. Upon further inspection, copepods were the most abundant animals in the dataset. Among groups expected to have meroplankton, molluscs were the most species-rich, followed by annelids. Ascidians, bryozoans, and hydrozoans, which are typically dominant in fouling communities, were absent or scarce.

Some species tagged as possible NIS include:

Species	**Taxon**	**Biogeographic pattern**
Anchoa mitchilli	Actinopterygii (Bay Anchovy)	Northwest Atlantic, Gulf of Mexico
Paralichthys dentatus	Actinopterygii (Summer Flounder)	Northwest Atlantic
Micromonas pusilla	Chlorophyta (Geen algae)	Europe
Americamysis bigelowi	Crustacea (Mysid)	Northwest Atlantic, Gulf of Mexico
Melosira nummuloides	Diatom	North Atlantic, Gulf of Mexico
Navicula ramosissima	Diatom	Europe, New Zealand
Podosira stelligera	Diatom	Mostly Northeast Atlantic
Thoracosphaera heimii	Dinoflagellate	South Atlantic, Mediterranean, Indian Ocean
Tectura testudinalis	Gastropod (limpet, synonym = *Testudinalia testudinalis*)	North Atlantic, Baltic Sea
Flabellina verrucosa	Gastropod (Nudibranch, synonym = *Coryphella verrucosa*)	North Atlantic
Aeolidea papillosa	Gastropod (Nudibranch)	North and West Atlantic, Baltic Sea, a few records in Puget Sound or Alaska
Onchidoris bilamellata	Gastropod (Nudibranch)	North Atlantic, NE Pacific
Alderia modesta	Gastropod (Saccoglosssa)	North Atlantic
Attheya longicornis	Ochrophyta (Brown algae)	North Atlantic, Baltic Sea
Hincksia granulosa	Ochrophyta (Brown algae)	West Atlantic, Baltic Sea
Laminaricolax aecidioloides	Ochrophyta (Brown algae)	West Atlantic, Mediterranean
Alitta succinea	Polychaete (Nereidae)	North Atlantic

Species with no data in OBIS were not evaluated (Appendix B).

II. Comparisons across day and night

Alpha diversity
To examine differences in zooplankton communities across day and night, a subset of 24 samples, with 77 OTUs, and 3,054,953 reads was created (i.e., DVN dataset; Table 2). Using the Chao1 diversity metric, alpha diversity appeared highest in May at the Valdez Marine Terminal

(no samples were collected in May at the Container Terminal; Figure 7). There did not appear to be any differences in alpha diversity across day and night.

Beta diversity
The NMDS plots generated with the DVN dataset indicated that the community composition in the samples collected at day and night in both May and June did not appear different, as the ellipses of samples collected during the day and night clearly overlapped (Figure 8). The PERMANOVA indicated that community composition in day versus night samples were not significantly different ($p = 0.303$, DVN; Table 3).

III. Comparisons across tidal cycle

Alpha diversity
To examine differences in zooplankton communities across the tidal cycle, a subset of 45 samples, with 130 OTUs, and 4,168,976 reads was created (i.e., TIDE dataset; Table 2). Using the Chao1 diversity metric, alpha diversity appeared to be relatively similar across tides within a month, but oscillated across months (Figure 9).

Beta diversity
The NMDS plots generated with the TIDE dataset indicated that the community composition in the samples collected across the tidal cycle within a month were not different, as the ellipses of samples collected during the different phases of the tide overlapped (Figure 10). The PERMANOVA results indicated that while tidal cycle did not significantly impact community composition ($p = 0.771$, TIDE-Tide; Table 3), month sampled did ($p = 0.001$, TIDE-Month; Table 3). The Tukey test for the TIDE-Month dataset indicated that there was a dispersion event ($p < 0.0001$; TIDE-Month; Table 4), which was evident given the spread of samples in the NMDS plots. Combining the output from the NMDS plot and these results, it appeared that both month and dispersion have significant effects.

DISCUSSION

Expanded sampling, compared to our previous studies in Port Valdez, allowed evaluation of sources of variation in plankton communities. However, we note that species accumulation curves (Figure 3) indicated that the number of samples and sequencing depth achieved did not fully capture the species diversity present in Port Valdez. Greater and deeper sampling will recover more rare species, though these may not be taxonomically assignable (if they lack representation in sequence databases) and may not be animals. Thus, our discussion is limited to species that could be identified.

Taxa
The majority of sequences in the zooplankton samples were assigned to copepods (Figure 6). Although sequence abundance is not a straightforward proxy of organismal abundance, this is expected as copepods are typically the most abundant animal taxon in marine plankton. Unfortunately, the preponderance of copepod sequences dilutes those belonging to more rare species, potentially reducing our ability to reconstruct benthic community composition. Many

molluscs were observed (Appendix A), while other taxa expected in nearshore Alaska were few or absent, such as anemones, flatworms, nemerteans, sponges, sipunculids, crabs, and shrimp. It is possible their absence is due to washout or dilution of their sequences by the sequences of the more abundant taxa. Another factor may be a greater number of brooding species in high latitude marine communities, compared to more equatorial sites, a pattern known as Thorsen's rule. In other words, fewer meroplankton might exist in Port Valdez compared to coastal waters in the contiguous Pacific states of the USA if those benthic taxa use other modes of reproduction.

As in previous years, important taxa that are usually abundant in fouling communities were not seen, including ascidians, bryozoan, and hydrozoans. Additionally, Ruiz et al. (2017) also found few ascidians and hydrozoans in a morphological assessment of fouling communities in Port Valdez (Table 5). Bryozoans were more represented in the morphological survey than in our plankton samples. These taxa often have short-lived larvae and may not disperse far from adult populations. Too, fouling communities likely occupy a small fraction of the total benthic habitat in Port Valdez. Thus, their relative scarcity in zooplankton samples and lack of abundance in morphological samples may reflect the relative size of adult populations compared to those in soft sediments and rocky shores. While we endeavored to sample physically closer to the fouling communities (through dockside sampling) more likely to contain NIS than in previous years, plankton sampling still missed many species found in the 2017 morphological survey. On the other hand, the total number of species detected and identified was much greater using the plankton metabarcoding approach as compared to the morphological only surveys. Some hybrid approach might be optimal for future detection of NIS.

Nonindigenous species (NIS)

We examined maps of global species distributions for all identified species with records in OBIS. Native species can fall into one of four patterns: 1) endemic to the temperate Northeastern Pacific (e.g., California to Alaska); 2) endemic to the North Pacific; 3) global at high northern latitudes (circumboreal); or 4) truly cosmopolitan (which may be more likely in holoplankton species). However, many recent genetic studies have shown that very widespread species (i.e., those in groups 3 and 4) are often species complexes. For those "species", in-depth phylogenetic studies are needed to distinguish between invasion and species complexes. Further, incomplete geographic sampling might misleadingly suggest sudden occurrence in Alaska, whereas Alaska records may simply be sparse in OBIS. Conversely, misidentified specimens in the OBIS database can confuse the actual geographic distribution of a species. Given these caveats, 17 species stood out for further investigation as potential NIS or new members of a cryptic species in Port Valdez (for an example see Figure 12). Absent from Port Valdez were common invasive species that would be expected from sources in California, Oregon, or Washington, such as *Mytilus galloprovincialis*, *Botrylloides violaceus*, and *Watersipora subatra*.

Scales of variation

The primary aim of this study was to determine significant sources of variation in plankton community composition and, in particular, meroplankton communities. The primary source of variation was the transition from spring to summer conditions (Tables 3 and 4, Figures 5 and 10), presumably reflecting temporal patterns as populations respond to seasonal increases in primary production. Interestingly, the significant effect of tidal conditions across months (Figure 9) may

suggest that on some dates, the efflux from the mudflat upshore from the Container Terminal contains a significantly different plankton community than the bay water rising at flood tide.

We saw no evidence of variation due to day or night (Table 3 and 4, Figures 7 and 8). In retrospect, vertical tows will sample across depths, so our design could not detect vertical phototaxis (the original design included depth stratified sampling).

We compared species lists from five years of metabarcoding surveys and found 155 of 258 identified species to occur in one year only and only nine found in all five years (Figure 11). Sampling effort varied from year to year, so a statistical comparison of yearly differences is difficult. Yet it appears that variation in species detection across years is as strong or stronger as within-year seasonal variation.

Summary and recommendations

1) Present data suggest that sequencing depth has been insufficient to fully capture animal OTU diversity in Port Valdez. A few species dominate the samples. Given this, a seasonal shift was nonetheless observed. For species detection, focusing on increased sequencing depth or molecular strategies to suppress dominant species might be considered.
2) Variation in community composition was primarily attributed to date of sampling and not day/night or tidal stage. A spring to summer shift was noticed, consistent with known plankton dynamics in Port Valdez. Increased replication of summer sampling might be considered to increase potential detection of meroplankton.
3) Taxa that are hallmarks of fouling communities were underrepresented and some species found by morphological surveys did not appear in our results. But the reverse is also true: metabarcoding found and identified many more species in Port Valdez than traditional visual surveys by a large margin. A hybrid strategy in which fouling communities are directly sampled and analyzed by metabarcoding might be considered. Additionally, waterborne eDNA, instead of plankton, might be collected from within the fouling community.

REFERENCES

1.
Altschul, S.F. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215, 403-410.
2.
Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J. & Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*, 13, 581-583.
3.
Carlton, J.T. & Geller, J. (1993). Ecological roulette: the global transport of nonindigenous marine organisms. *Science*, 261, 78-82.
4.
Chamberlain, S. (2018). worms: World Register of Marine Species (WoRMS) Client.

5.
Geller, J., Meyer, C., Parker, M. & Hawk, H. (2013). Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Mol Ecol Resour*, 13, 851-861.
6.
McMurdie, P.J. & Holmes, S. (2013). phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*, 8, e61217.
7.
OBIS (2002). Ocean Biodiversity Information System. Available at: www.obis.gov Last accessed March 25, 2022.
8.
Okasanen, J., Guillaume Blanchet, F., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B. *et al.* (2014). Package 'vegan': Community Ecology Package.
9.
Team, R.C. (2020). R: Foundation for statistical computing. Vienna, Austria.

Table 1. Sampling scheme used to assess the factors most likely influencing zooplankton communities including 1) time, 2) location, 3) daylight, and 4) tidal cycle. Due to access issues at the Valdez Marine Terminal, the fewest samples were collected from this location. Tidal cycle sampling was conducted at the Container Dock only. Day and night sampling was conducted at the Container Dock and the Small Boat Harbor (as referred to as Valdez Harbor).

	Onset of spawning				Peak Spawing and Settlement								Diminishing settlement									
	April				May				June				July				August				September	
	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12	Week 13	Week 14	Week 15	Week 16	Week 17	Week 18	Week 19	Week 20	Week 21	TOTAL
Site 1: Valdez Marine Terminal																						
Days of sampling	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1	6
Replicates at 5 meters Day	3	3	3	0	3	3	3	0	3	3	3	0	3	3	3	0	3	3	3	0	3	48
Replicates at 5 meters Night	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Site 2: Container Dock																						
Days of sampling	1	1	1	1	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	29
Replicates at 5 meters Day	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	63
Replicates at 5 meters Night	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Site 3: Small Boat Harbor																						
Days of sampling	1	1	1	1	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	29
Replicates at 5 meters Day	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	63
Days of sampling nights					1		1		1		1											4
Replicates at 5 meters Night	0	0	0	0	3	0	3	0	3	0	3	0	0	0	0	0	0	0	0	0	0	12
Site 2: Container Dock																						
Days of sampling					0	1	0	1	0	1	0	1										4
Tidal cycle - Slack					0	3	0	3	0	3	0	3										12
Tidal cycle - Ebb					0	3	0	3	0	3	0	3										12
Weekly sample size	**6**	**9**	**6**	**6**	**15**	**21**	**15**	**18**	**15**	**21**	**15**	**18**	**6**	**9**	**6**	**6**	**6**	**9**	**6**	**6**	**9**	**228**

Table 2. The number of samples, OTUs, and reads across each of the datasets analyzed in this report. The datasets were parsed so that statistical analyses could be conducted on an equal number of samples per factor. These included 1) Total (all samples with all OTUs), 2) Total_Animals (all samples with all OTUs identified as animals), 3) CON (all samples from the Container Dock), 4) VDZ (all samples from the Valdez Harbor), 5) VMT (all samples from the Valdez Marine Terminal), 6) DVN (selected samples for the day versus night comparison), 7) All3 (selected samples for comparison across locations), 8) TIDE (selected samples for comparison across tides). All the parsed datasets were parsed from the Total_Animals dataset, so only animals are included in analyses.

Dataset	**Sample #**	**OTU#**	**Read #**
Total	222	1,257	31,206,244
Total_Animals	222	195	24,447,209
CON	114	157	12,220,171
VDZ	99	138	10,301,018
VMT	18	74	1,926,020
DVN	24	77	3,054,953
All3	54	94	5,581,255
Tide	45	130	4,168,976

Table 3. The results of the PERMANOVA tests conducted in the vegan package in R for each of the three datasets. For the All3 datasets, the significance of both location and month were tested and the results of both are shown. Statistical significance was based on a p-value ≤ 0.05.

	DVN	**All3-Location**	**All3-Month**	**Tide-Tide**	**Tide-Month**
Degrees of freedom	1	2	5	2	2
Sums of Squares	0.2225	0.6565	5.8942	0.4118	3.7471
Mean Squares	0.22249	0.32824	1.17885	0.20589	1.87356
F. Model	1.1207	1.3432	7.8316	0.76563	9.8864
R2	0.04847	0.05004	0.44928	0.03518	0.32009
Pr(>F)	0.303	0.122	**0.001*** **	0.771	**0.001*** **

Table 4. As a follow-up to the result of the PERMANOVA tests, we also conducted Tukey tests in the vegan package in R. For the All3 dataset, the significance of both location and month were tested and both results are shown. For the TIDE dataset, the significance of both tide and month were tested and both results are shown. Statistical significance was based on a p-value ≤ 0.05.

	DVN	**All3-Location**	**All3-Month**	**Tide-Tide**	**Tide-Month**
Degrees of freedom	1	2	5	2	2
Sums of Squares	0	0.009	1.4005	0.22865	1.98
Mean Squares	0	0.0045	0.280097	0.114325	0.9878
F value	0	0.0481	10.334	1.5446	69.651
Pr(>F)	0.9997	0.9531	**<0.0001*** **	0.2253	**<0.0001*** **

Table 5. Results of 2016 morphological survey conducted by the Marine Invasions Research Laboratory at the Smithsonian Environmental Research Center (see Table 2 from Ruiz et al., 2017). Taxonomic overlap between zooplankton samples from this study and benthic samples from their study are shown in bold.

Anthozoa	Anemone sp (1 or 2 spp)
Bryozoa	*Alcyonidium* sp
	Bugula pacifica
	Callopora sp
	Celleporella hyalina
	Crissidae sp
	Dendrobeania sp
	Fenestrulina delicia
	Membranipora villosa
	Primaverans sp
	Rhynchozoon sp
	Tubulipora cf pacifica
Crustacea	***Balanus* sp**
Echinodermata	*Pisaster* sp
Hydrozoa	cf *Obelia* sp
	cf *Clytia* sp
Molluscs	***Dendronotus* sp**
	Dorid Nudibranch
	Hermissenda crassicornis
	cf *Pododesmus* sp
	Hiatella arctica
	Mytilus cf trossulus
	scallop
	slipper limpet
Polychaeta	*Crucigera zygophora*
	Dorvillaidae
	Nereidae
	Pseudochitinopoma occidentalis
	Serpula sp
	Spirorbidae sp 1
	Spirorbidae sp 2
Porifera	Unidentified sponge
	cf *Halichondria* sp
	Fiberglass sponge
Tunicata	*Corella inflata*
	cf *Halocynthia* sp

Figure 1. Alpha diversity metrics using the Chao1 diversity metric of animal OTU richness across the three sampled sites using the CON, VDZ, and VMT datasets across months (Total_Animal dataset). Note that all samples within the month are pooled for this analysis.

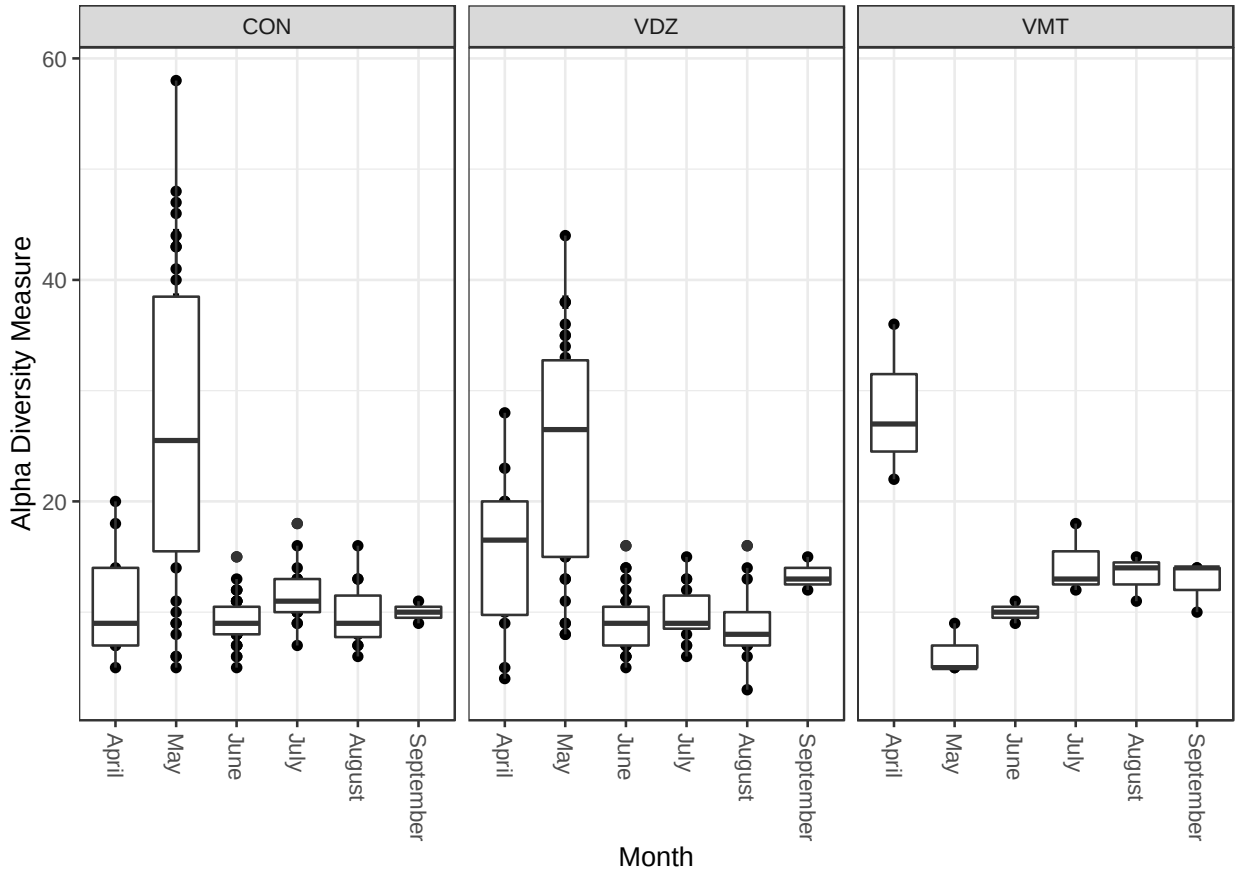


Figure 2. Alpha diversity metrics using the Chao1 diversity metric of animal OTU richness across the three sampled sites using the CON, VDZ, and VMT datasets across months (All3 dataset). Note that all samples within the month are pooled for this analysis.

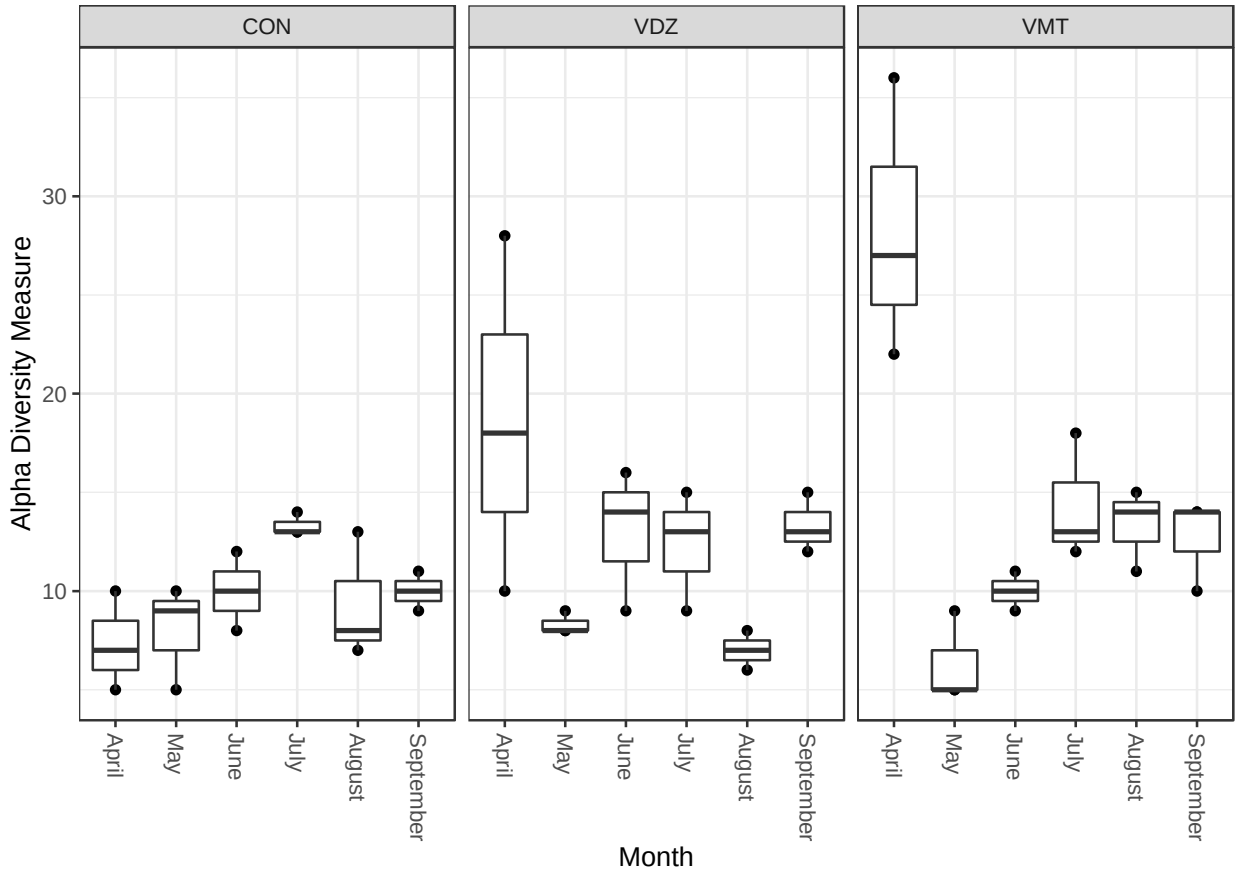


Figure 3. Animal OTU accumulation curves created in the vegan package in R for all the sequence data combined (Total_Animals: A), then parsed by location, the CON (B), the VDZ (C), and the VMT (D) datasets. Note the difference in the values of the x and y axes of (A) compared to the other graphs.

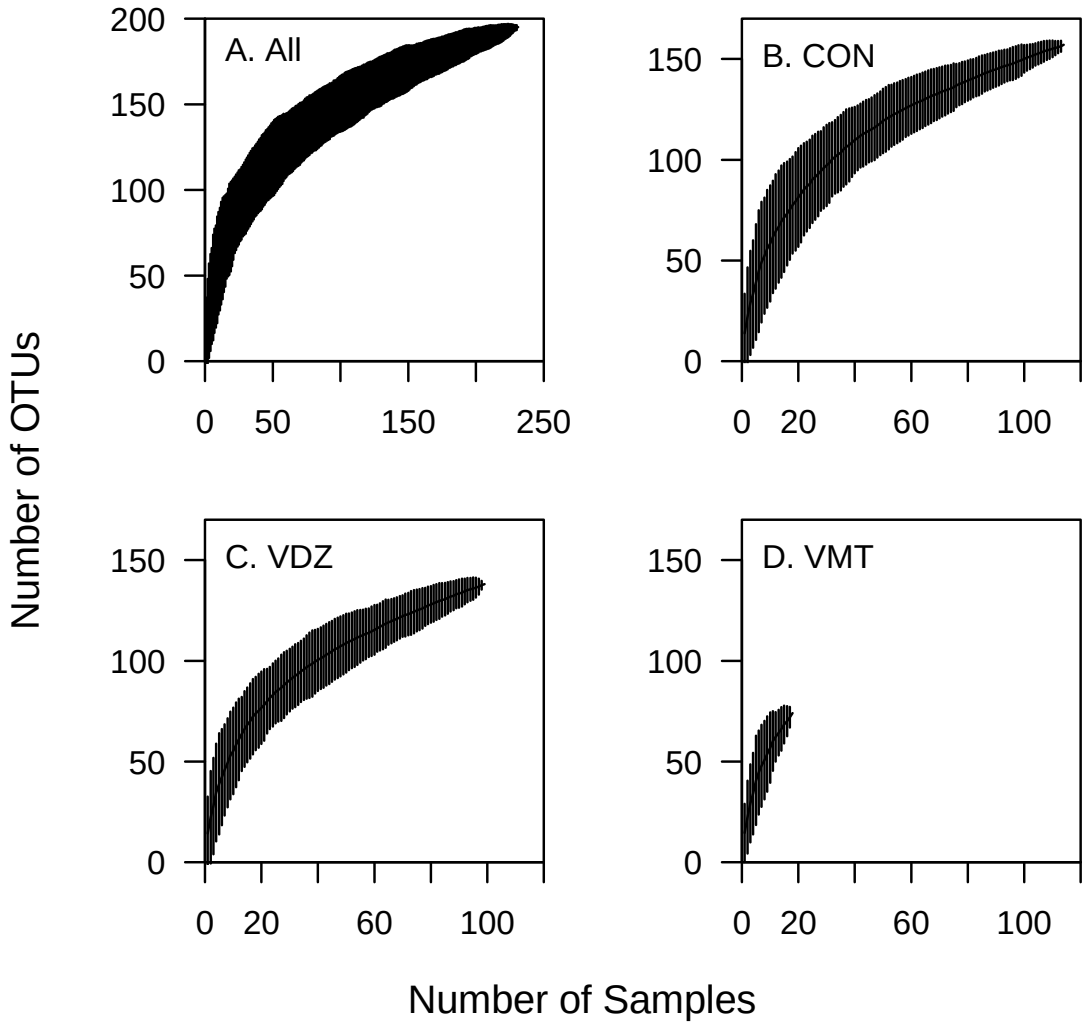


Figure 4. Principal Coordinates Analysis (PCoA) plots for CON (A), VDZ (B), and VMT (C) using all samples collected from each location. Coloring corresponds to the month in which samples were collected. Ellipses could not be calculated for the VMT dataset due to the small number of samples collected.

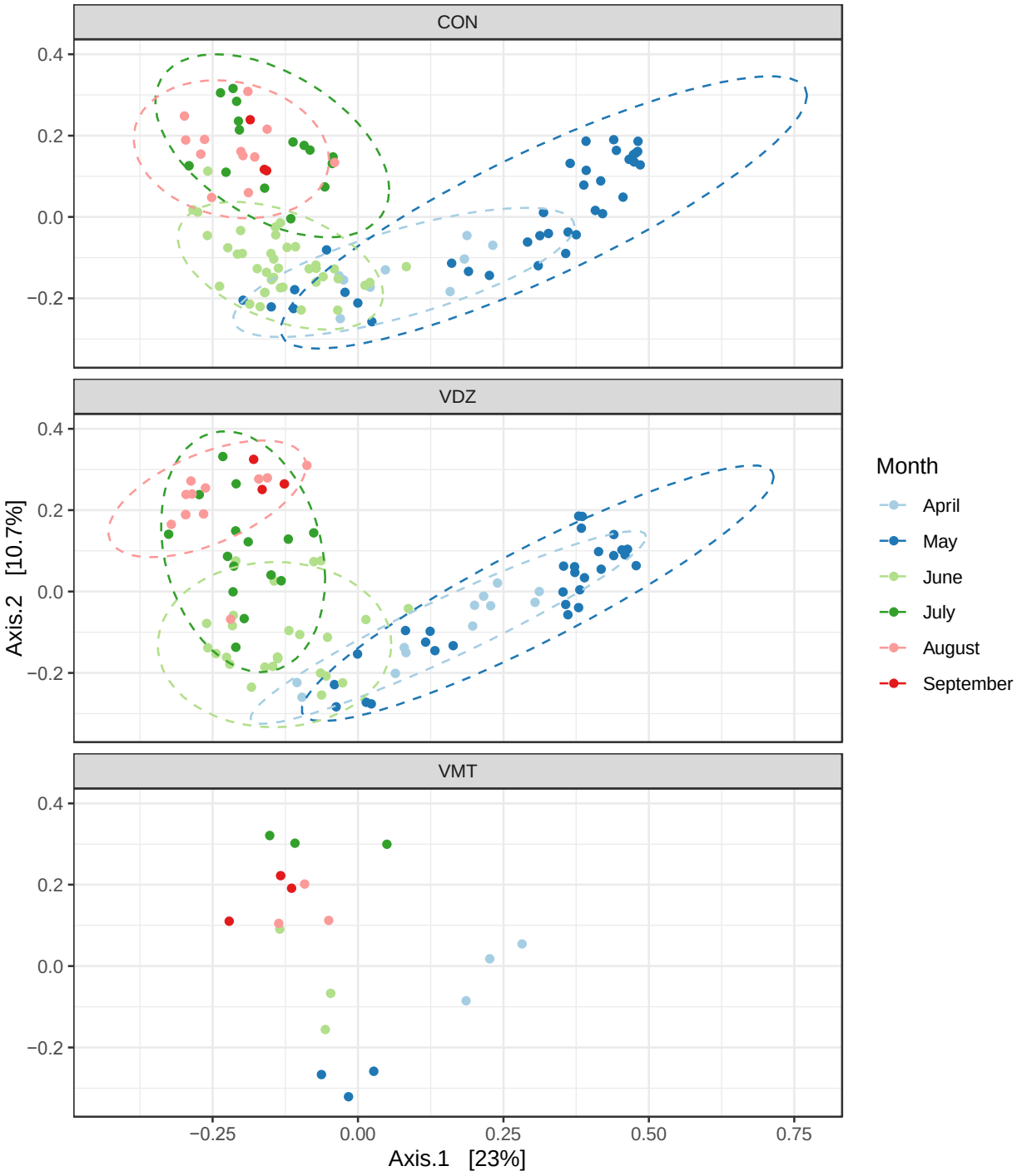


Figure 5. Non-metric multidimensional scaling (NMDS) plot for the All3 dataset to assess the impact of month on the community composition of the samples. Coloring corresponds to the month in which samples were collected.

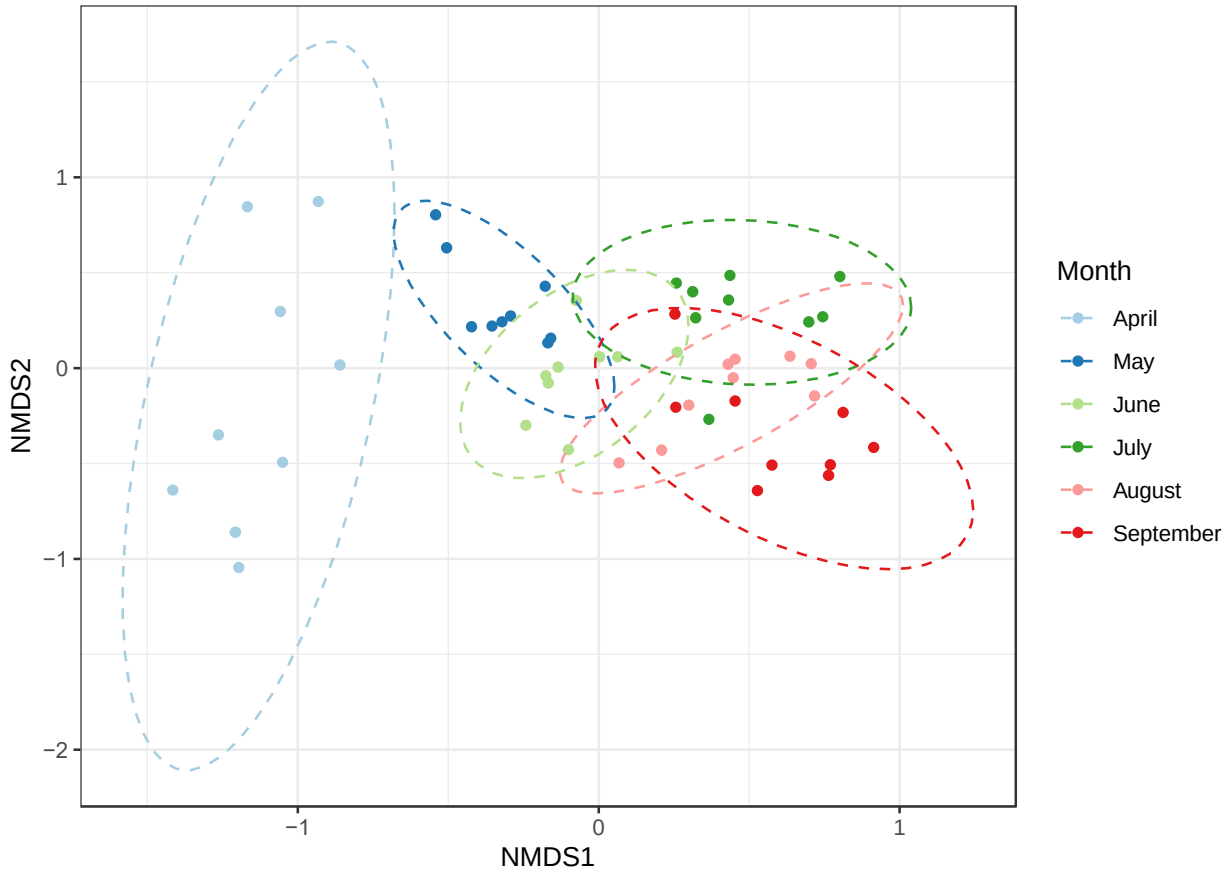


Figure 6. The sequence abundance of each phylum collected at each sampling location across all the months where samples were obtained. This graph was generated using all available samples (i.e., the Total_Animals dataset).

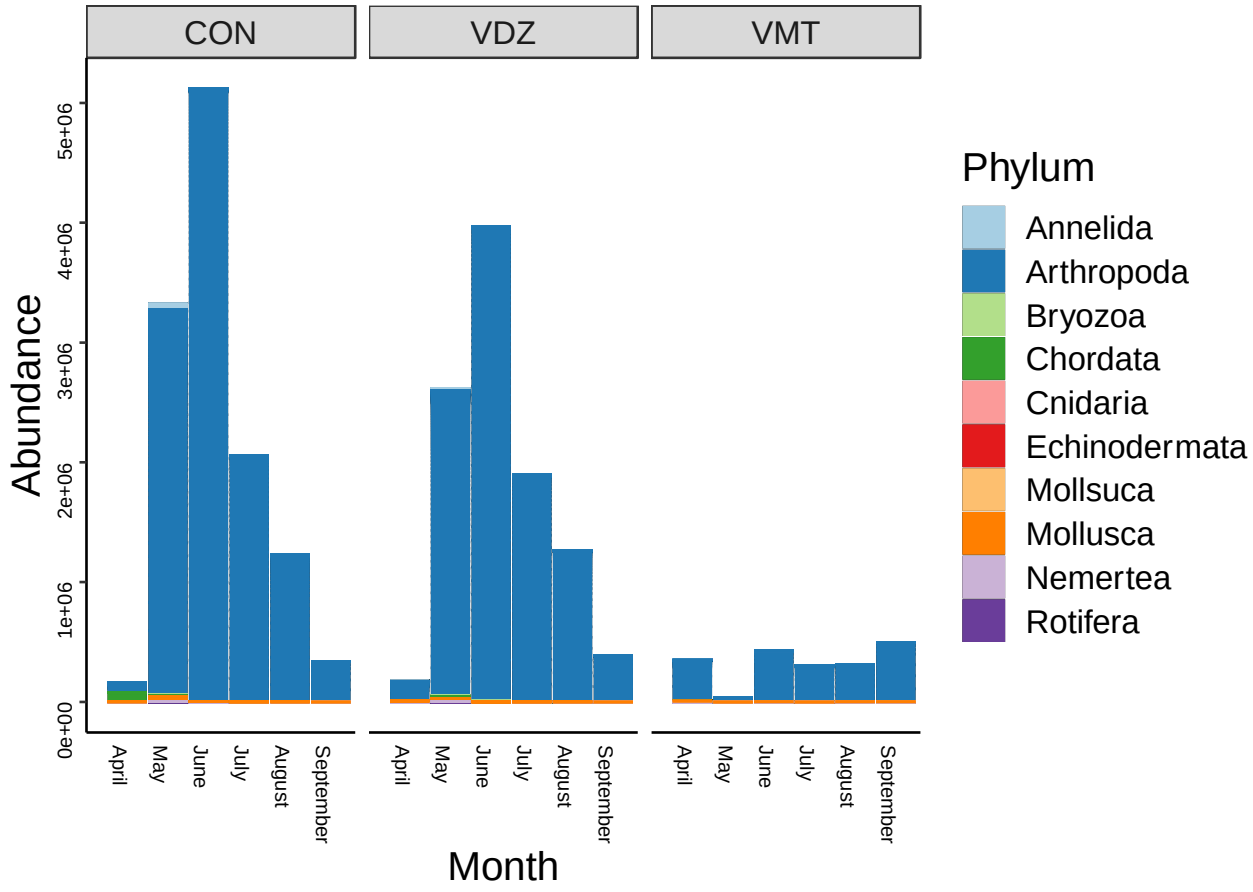


Figure 7. Alpha diversity metrics using the Chao1 diversity metric of animal OTU richness during the day and night at both CON and VDZ (DVN dataset). Colors shown indicate the months in which the samples were collected, either May or June.

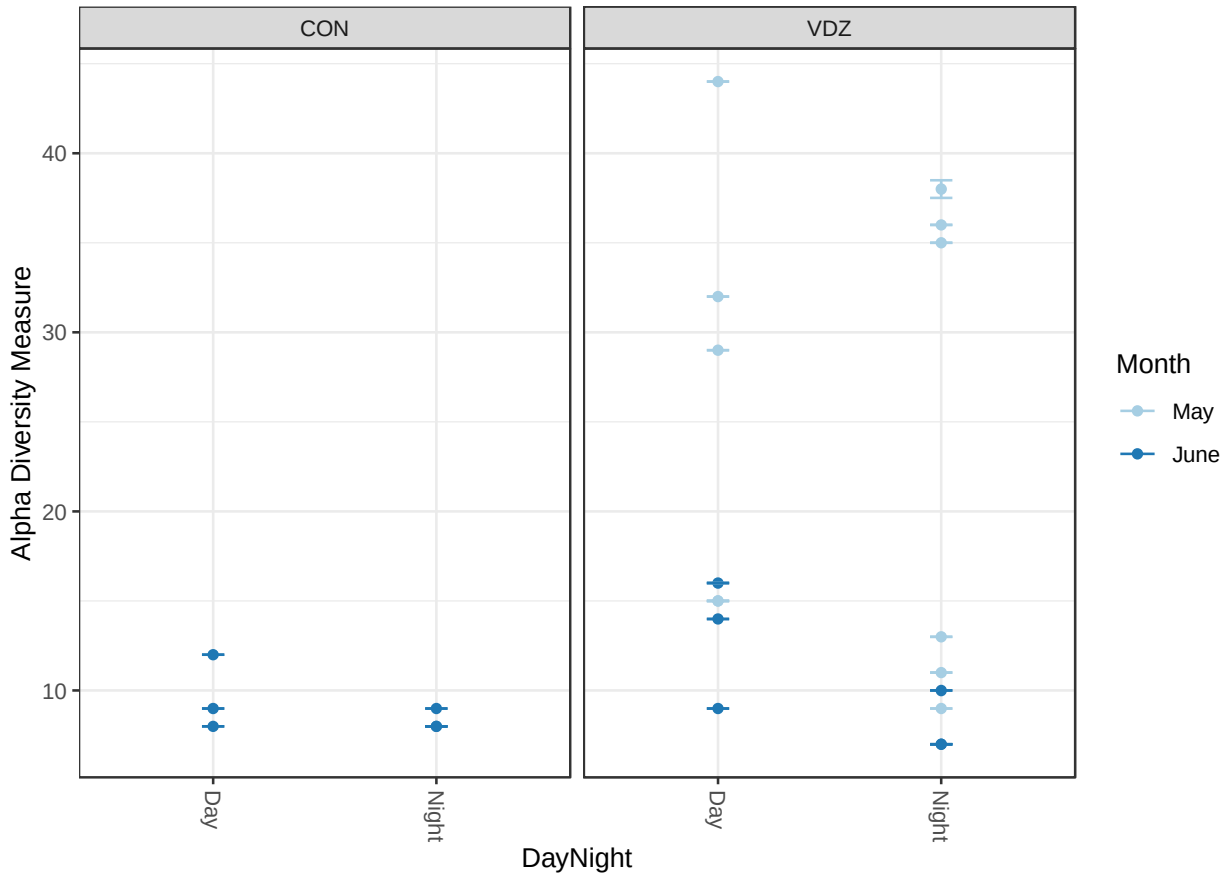


Figure 8. Non-metric multidimensional scaling (NMDS) plots for the DVN dataset to assess the impact of sampling at day versus night across the two months on the community composition of the samples. Coloring corresponds to the timing of the sampling.

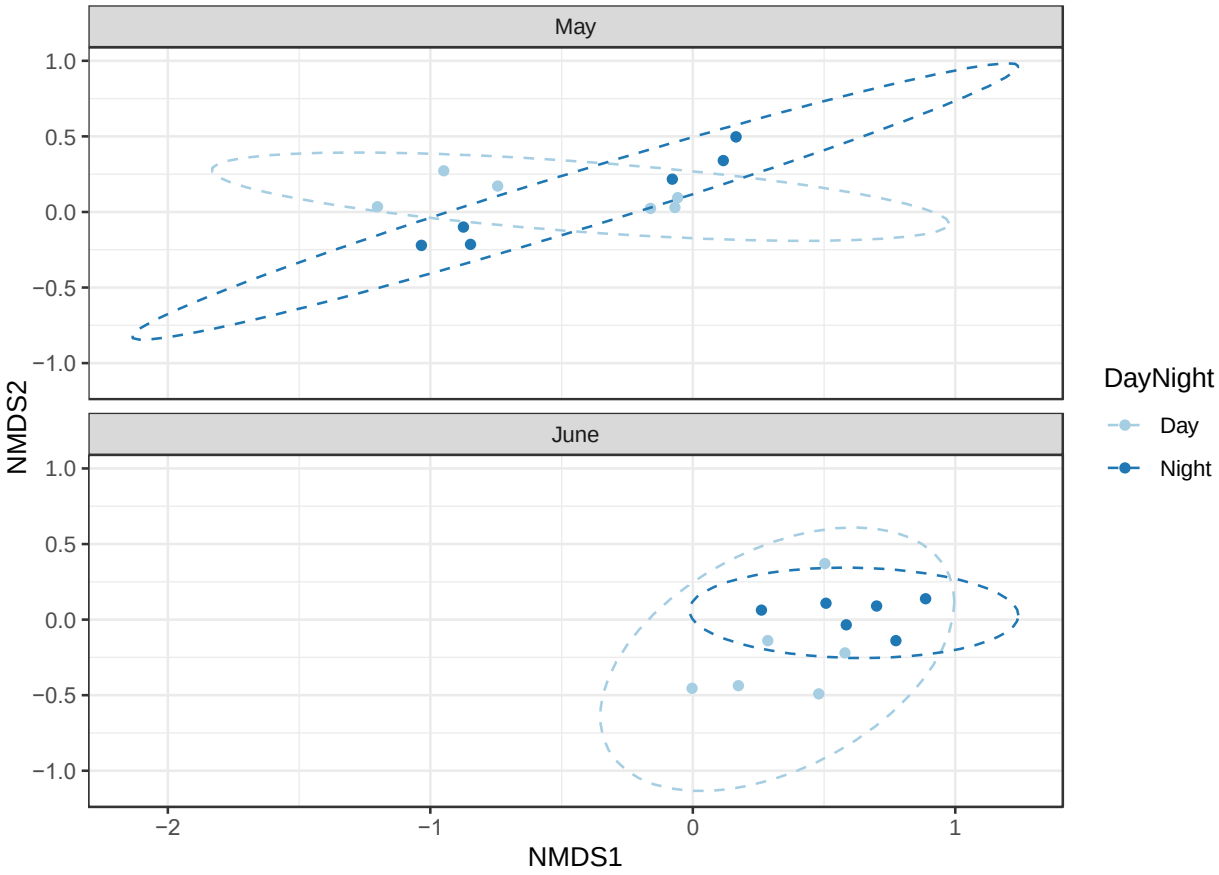


Figure 9. Alpha diversity metrics using the Chao1 diversity metric of animal OTU richness across the tidal cycle by month where samples were obtained (TIDE dataset).

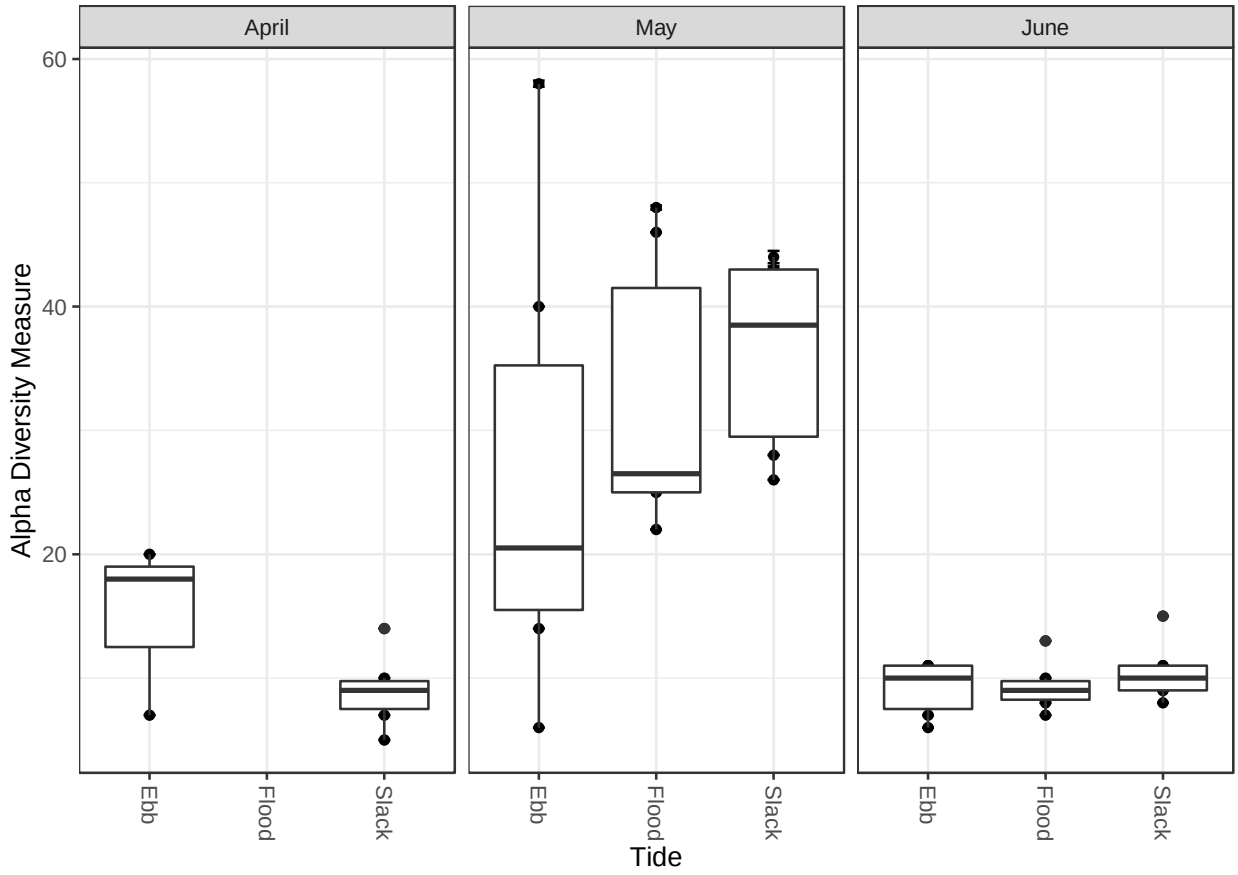


Figure 10. Non-metric multidimensional scaling (NMDS) plots for the TIDE dataset to assess the impact of the tidal cycle across months on the community composition of the samples. Coloring corresponds to the tidal cycle.

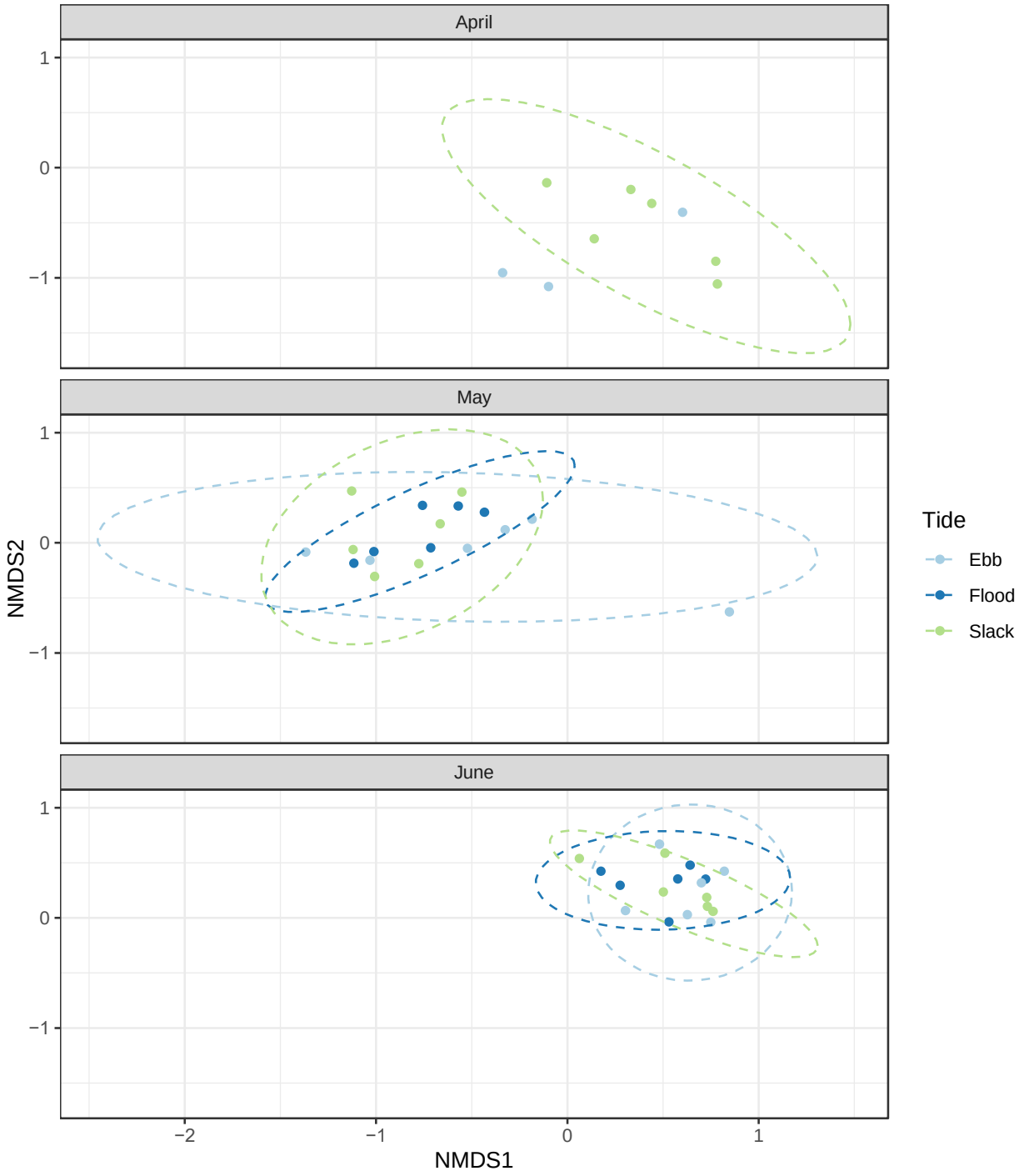


Figure 11. Frequency of occurrence of identified species (excluding additional taxa from September 2021 samples) in plankton samples from Port Valdez, 2016-2021. Note that stations and sample sizes varied by year.

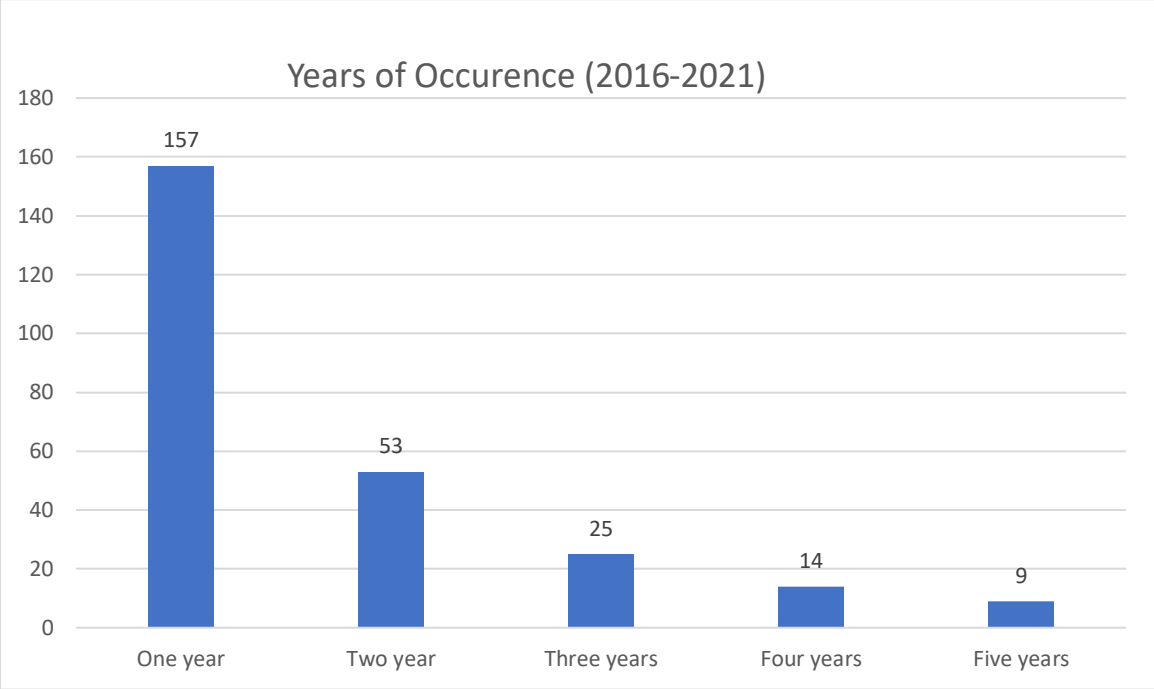
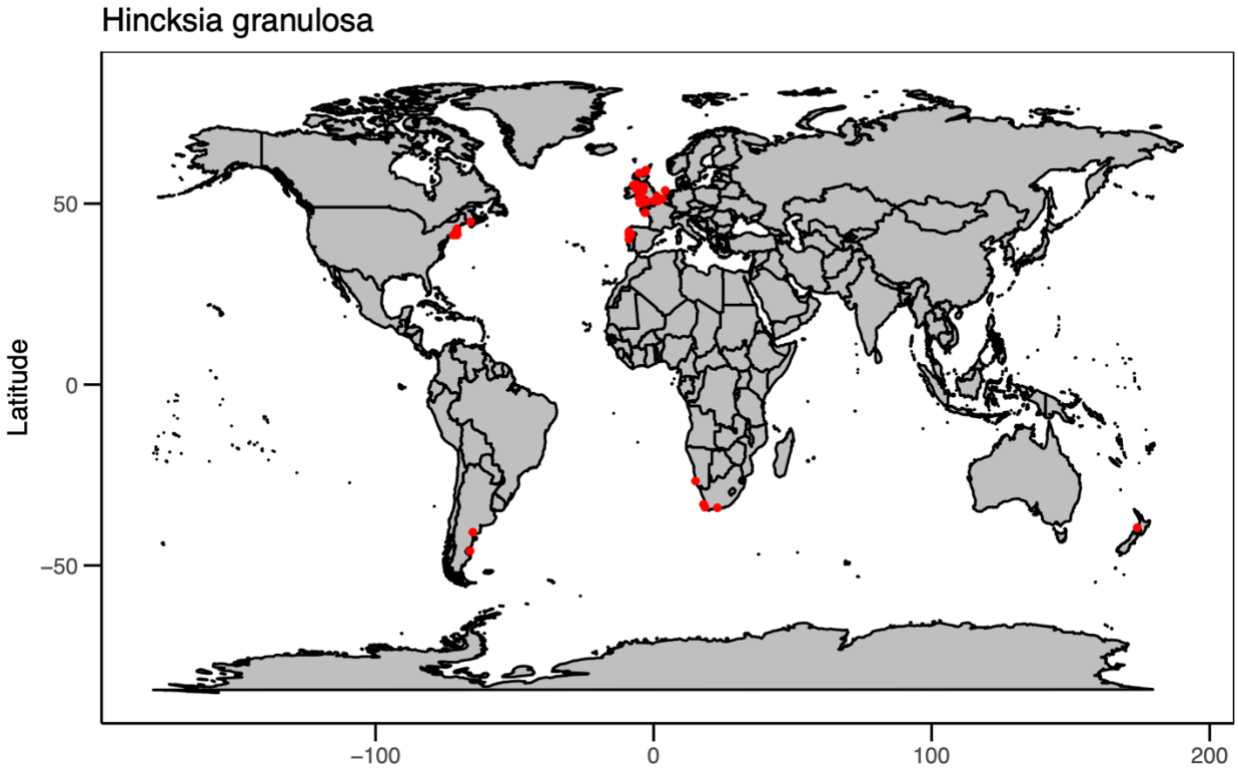


Figure 12. Global distribution of *Hincksia granulosa* from OBIS records suggesting the novel appearance in Port Valdez, Alaska. This brown alga was described in 1811 in Great Britain (https://www.algaebase.org/search/species/detail/?species_id=13016).



Appendix A. Animal species identified in Valdez plankton samples determined through BLAST against the MLML COI dataset and Genbank. Sequences with e-value of $1x10^{-30}$, 95% pairwise identify, and 90% coverage with database records were annotated to the lowest possible taxonomic level.

Kingdom	Phylum	Order	ScientificName
Animalia	Arthropoda	Calanoida	*Acartia hudsonica*
Animalia	Arthropoda	Calanoida	*Acartia longiremis*
Animalia	Mollusca	Nudibranchia	*Aeolidia libitinaria*
Animalia	Mollusca	Nudibranchia	*Aeolidia papillosa*
Animalia	Bryozoa	Ctenostomatida	*Alcyonidium polyoum*
Animalia	Mollusca	NA	*Alderia modesta*
Animalia	Annelida	Phyllodocida	*Alitta succinea*
Animalia	Arthropoda	Mysida	*Americamysis bigelowi*
Animalia	Chordata	Clupeiformes	*Anchoa mitchilli*
Animalia	Mollusca	Nudibranchia	*Apata pricei*
Animalia	Mollusca	Sacoglossa	*Aplysiopsis enteromorphae*
Animalia	Mollusca	NA	*Aplysiopsis enteromorphae*
Animalia	Annelida	Echiuroidea	*Arhynchite pugettensis*
Animalia	Arthropoda	Balanomorpha	*Balanus*
Animalia	Arthropoda	Balanomorpha	*Balanus crenatus*
Animalia	Arthropoda	Balanomorpha	*Balanus glandula*
Animalia	Arthropoda	Balanomorpha	*Balanus rostratus*
Animalia	Annelida	Phyllodocida	*Bipalponephtys neotena*
Animalia	Cnidaria	Anthoathecata	*Bougainvillia superciliaris*
Animalia	Arthropoda	Calanoida	Calanoida
Animalia	Arthropoda	Calanoida	*Calanus marshallae*
Animalia	Arthropoda	Calanoida	*Calanus pacificus*
Animalia	Nemertea	Monostilifera	*Carcinonemertes epialti*
Animalia	Arthropoda	Calanoida	*Centropages abdominalis*
Animalia	Nemertea	Heteronemertea	*Cerebratulus*
Animalia	Annelida	Sabellida	*Chone*
Animalia	Cnidaria	Semaeostomeae	*Chrysaora melanaster*
Animalia	Chordata	Pleuronectiformes	*Citharichthys stigmaeus*
Animalia	Chordata	Perciformes	*Clinocottus acuticeps*
Animalia	Mollusca	Pteropoda	*Clione*
Animalia	Chordata	Clupeiformes	*Clupea pallasii*
Animalia	Cnidaria	Leptothecata	*Clytia gregaria*
Animalia	Mollusca	Nudibranchia	*Corambe steinbergae*
Animalia	Mollusca	Littorinimorpha	*Crepipatella lingulata*
Animalia	Arthropoda	Cyclopoida	*Cyclops columbianus*

Animalia	Mollusca	Nudibranchia	*Dendronotus albopunctatus*
Animalia	Mollusca	Nudibranchia	*Dendronotus albus*
Animalia	Mollusca	Nudibranchia	*Dendronotus subramosus*
Animalia	Mollusca	Nudibranchia	*Dendronotus venustus*
Animalia	Mollusca	Nudibranchia	*Doris montereyensis*
Animalia	Annelida	Eunicida	Dorvilleidae
Animalia	Annelida	Phyllodocida	*Eteone*
Animalia	Annelida	Phyllodocida	*Eteone longa*
Animalia	Arthropoda	Calanoida	*Eucalanus bungii*
Animalia	Annelida	Phyllodocida	*Eulalia quadrioculata*
Animalia	Annelida	Phyllodocida	*Eulalia viridis*
Animalia	Annelida	Phyllodocida	*Eunoe*
Animalia	Arthropoda	Euphausiacea	*Euphausia pacifica*
Animalia	Arthropoda	Cyclopoida	*Euryte*
Animalia	Arthropoda	Onychopoda	*Evadne nordmanni*
Animalia	Echinodermata	Forcipulatida	*Evasterias troschelii*
Animalia	Mollusca	Nudibranchia	*Flabellina trilineata*
Animalia	Mollusca	Nudibranchia	*Flabellina verrucosa*
Animalia	Annelida	NA	*Galathowenia oculata*
Animalia	Annelida	Phyllodocida	*Gattyana cirrhosa*
Animalia	Annelida	Phyllodocida	*Glycera nana*
Animalia	Cnidaria	Anthoathecata	*Halitholus*
Animalia	Annelida	Phyllodocida	*Harmothoe*
Animalia	Annelida	Phyllodocida	*Harmothoe extenuata*
Animalia	Arthropoda	Harpacticoida	Harpacticoida
Animalia	Arthropoda	Harpacticoida	*Harpacticus uniremis*
Animalia	Mollusca	Adapedonta	*Hiatella*
Animalia	Mollusca	Nudibranchia	*Himatina trophina*
Animalia	Arthropoda	Decapoda	Hippolytidae
Animalia	Mollusca	Nudibranchia	*Janolus fuscus*
Animalia	Mollusca	Galeommatida	*Kellia suborbicularis*
Animalia	Mollusca	Nudibranchia	*Knoutsodonta jannae*
Animalia	Mollusca	Littorinimorpha	*Lacuna vincta*
Animalia	Annelida	Spionida	*Laonice*
Animalia	Annelida	Spionida	*Laonice cirrata*
Animalia	Mollusca	Venerida	*Leukoma staminea*
Animalia	Chordata	Pleuronectiformes	*Limanda aspera*
Animalia	Mollusca	Cardiida	*Limecola balthica*
Animalia	Nemertea	Heteronemertea	*Lineus*

Animalia	Nemertea	Heteronemertea	*Lineus flavescens*
Animalia	Mollusca	Cardiida	*Macoma calcarea*
Animalia	Nemertea	Heteronemertea	*Maculaura aquilonia*
Animalia	Nemertea	Heteronemertea	*Maculaura cerebrosa*
Animalia	Annelida	NA	*Magelona*
Animalia	Mollusca	Trochida	*Margarites pupillus*
Animalia	Mollusca	Cephalaspidea	*Melanochlamys diomedea*
Animalia	Cnidaria	Leptothecata	*Melicertum octocostatum*
Animalia	Bryozoa	Cheilostomatida	*Membranipora villosa*
Animalia	Arthropoda	Harpacticoida	*Mesochra*
Animalia	Arthropoda	Decapoda	*Metacarcinus gracilis*
Animalia	Arthropoda	Calanoida	*Metridia pacifica*
Animalia	Arthropoda	Calanoida	*Microcalanus*
Animalia	Annelida	Phyllodocida	*Micronereis nanaimoensis*
Animalia	Mollusca	Mytilida	*Mytilus trossulus*
Animalia	Mollusca	Neogastropoda	*Nassarius mendicus*
Animalia	Arthropoda	Calanoida	*Neocalanus flemingeri*
Animalia	Arthropoda	Calanoida	*Neocalanus plumchrus*
Animalia	Mollsuca	NA	*Odostomia*
Animalia	Arthropoda	Cyclopoida	*Oithona similis*
Animalia	Mollusca	Sacoglossa	*Olea hansineensis*
Animalia	Mollusca	Nudibranchia	*Onchidoris bilamellata*
Animalia	Annelida	NA	*Ophelia*
Animalia	Echinodermata	Ophiurida	*Ophiura sarsii*
Animalia	Arthropoda	Diptera	Orthocladiinae
Animalia	Annelida	Eunicida	*Palpiphitime lipovskyae*
Animalia	Arthropoda	Calanoida	*Paracalanus*
Animalia	Arthropoda	Calanoida	*Paraeuchaeta elongata*
Animalia	Chordata	Pleuronectiformes	*Paralichthys dentatus*
Animalia	Nemertea	Monostilifera	*Paranemertes californica*
Animalia	Arthropoda	Calanoida	*Pareucalanus attenuatus*
Animalia	Annelida	Terebellida	*Pectinaria granulata*
Animalia	Arthropoda	Harpacticoida	Peltidiidae
Animalia	Annelida	unknown	*Pharyngocirrus uchidai*
Animalia	Annelida	Phyllodocida	*Pholoe*
Animalia	Annelida	Phyllodocida	*Pholoides asperus*
Animalia	Annelida	Phyllodocida	*Phyllodocidae*
Animalia	Arthropoda	Decapoda	Pinnotheridae
Animalia	Annelida	Terebellida	*Pista wui*
Animalia	Arthropoda	Onychopoda	*Pleopis polyphemoides*

Animalia	Annelida	Phyllodocida	*Podarkeopsis perkinsi*
Animalia	Arthropoda	Onychopoda	*Podon leuckartii*
Animalia	Arthropoda	Copepoda	*Poecilostomatoida*
Animalia	Annelida	NA	*Polygordius*
Animalia	Annelida	Phyllodocida	Polynoidae
Animalia	Nemertea	Monostilifera	*Poseidonemertes collaris*
Animalia	Annelida	Spionida	*Prionospio steenstrupi*
Animalia	Arthropoda	Diptera	*Psectrocladius limbatellus*
Animalia	Arthropoda	Calanoida	*Pseudocalanus*
Animalia	Arthropoda	Calanoida	*Pseudocalanus mimus*
Animalia	Arthropoda	Calanoida	*Pseudocalanus minutus*
Animalia	Arthropoda	Calanoida	*Pseudocalanus moultoni*
Animalia	Mollusca	Littorinimorpha	Ranellidae
Animalia	Annelida	Spionida	*Rhynchospio glutaea*
Animalia	Annelida	NA	Sabellariidae
Animalia	Annelida	unknown	*Saccocirrus*
Animalia	Mollusca	Venerida	*Saxidomus gigantea*
Animalia	Annelida	NA	*Scoloplos armiger*
Animalia	Arthropoda	Balanomorpha	*Semibalanus balanoides*
Animalia	Arthropoda	Balanomorpha	*Semibalanus cariosus*
Animalia	Arthropoda	Diptera	*Sphaerophoria philanthus*
Animalia	Annelida	Spionida	Spionidae
Animalia	Mollusca	NA	*Stiliger fuscovittatus*
Animalia	Cnidaria	Anthoathecata	*Stomotoca atra*
Animalia	Echinodermata	Camarodonta	*Strongylocentrotus droebachiensis*
Animalia	Rotifera	Ploima	*Synchaetidae*
Animalia	Annelida	Terebellida	*Terebellides stroemii*
Animalia	Mollusca	NA	*Testudinalia (Tectura) testudinalis*
Animalia	Arthropoda	Euphausiacea	*Thysanoessa inermis*
Animalia	Arthropoda	Euphausiacea	*Thysanoessa raschii*
Animalia	Arthropoda	Euphausiacea	*Thysanoessa spinifera*
Animalia	Arthropoda	Harpacticoida	*Tisbe*
Animalia	Mollusca	Littorinimorpha	*Trichotropis cancellata*