UNIVERSITY OF CALIFORNIA SAN DIEGO

Microbial Community Structure of the Sea Surface-Marine Boundary Layer

A thesis submitted in partial satisfaction of the
requirements for the degree
Master of Science

in

Oceanography

by

Matthew John Murray Heron

Committee in charge:

Professor Jeff Shovlowsky Bowman, Chair
Professor Andrew David Barton
Professor Lynn Monica Russell

2022

Thesis Approval Page

The thesis of Matthew Heron is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego
2022

Dedication

In recognition of the members of the Bowman Lab for their contributions to this degree ranging from laboratory work to critical feedback to emotional support.

Table of Contents

List of Figures

List of Tables

Abstract of The Thesis

Microbial Community Structure of the Sea Surface-Marine Boundary Layer

by

Matthew John Murray Heron

Master of Science in Oceanography

University of California San Diego, 2022

Professor Jeff Shovlowsky Bowman, Chair

Environmental processes aerosolize bacteria from the seawater and into the atmosphere. In the sea surface-atmospheric boundary layer (SSABL), two very distinct biomes can overlap. The SSABL's microbiota fluctuates between the two environments. Airborne microbiological habitats have few microbes, while the coastal sea surface has one of the most diverse microbial ecosystems.

Because of the ecosystem's variability, it is hard to paint the whole picture due to the region's significant daily irregularity, and addressing its criteria is a significant endeavor. We

faced challenges in refining our equipment for 21st-century microbiome analysis. Our design addressed equipment sterilization and low abundance in sample collection. Our approach aggregated the data over a month, delivering a glimpse of this dynamic ecological environment. This study aimed to characterize the SSABL community structure by identifying taxa likely to inhabit the interface and determining how similar these populations are to the parent marine community. Using 16S rRNA gene sequencing and flow cytometry, airborne samples were compared to marine samples. We found a link between meteorological conditions and the SSABL community by tracking air masses to the Scripps Pier with NOAA's HYSPLIT model.

Oceanic air masses are a component in influencing a site's community composition. Thus, we detected aerosolized marine taxa. We then describe factors influencing microbiome makeup using multiple regression and environmental variable ordination.

Introduction

The microbial community of the Sea Surface-Atmospheric Boundary Layer (SSABL) is part of an interface where microorganisms from the ocean and the atmosphere interact. Parameterization of this ecosystem is a topic of many studies, yet the community structure of this environment is still incompletely characterized. Marine microbial organisms are aerosolized into the atmosphere and participate in various chemical and physical processes in the atmosphere and the ocean (Bauer, 2003). Airborne microbes influence the ecosystems with which they interact. For example, microorganisms living at the interface may also play a significant role in global climate regulation by producing dimethyl sulfide, which is essential in forming clouds (Welsh, 2000). The SSABL is a small window between two ecosystems where the microbial community lives in a highly dynamic environment, and observing this domain is a formidable challenge. The two ecosystems are in constant flux and come into direct contact due to surface layer gas exchange (Engel et al., 2017). Some organisms travel long distances by moving passively with air masses (Mayol et al., 2017). Some bacteria can move between liquid and gas phases and survive at the interface; they can even germinate when deposited (Hu et al., 2017).

As a result, our team thought it would be an exciting opportunity to study the species composition and community structure of the SSABL in the coastal waters of La Jolla, California, at the Scripps Institution of Oceanography (SIO). Based on an ecological analysis, we set out to identify which species, and in what abundance, had the best chance of surviving in this interface.

To compare the microbial community of the seawater and atmospheric components of the SSABL, we used high-throughput next-generation sequencing of the 16S rRNA gene in conjunction with novel contemporary gene analysis techniques. We addressed microbial

community composition through phylogenetic analysis, microbial abundance analysis, and computing a Marine Similarity Coefficient (MSC) to determine how structurally comparable our airborne samples were to the exclusively marine samples. We assessed the origin of air mass at the pier using the backward trajectory model from NOAA's Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT). Finally, we used flow cytometry to quantify microbial abundances (FCM). We hypothesized that if the marine community interacts with the atmosphere, we should expect to find a significant marine microbial population in airborne samples, contributing a significant population to the SSABL.

Methods

Pier setup overview

Figure 1: Deployment of Bertin Coriolis Compact (CC).
The CC was Deployed from the North end of the Scripps Pier. Daily Samples were collected during a six and a half hour contious sampling peroid.

To address the significance of the marine microbial community's interaction with the atmospheric community, we collected atmospheric and seawater samples from the Ellen Browning Scripps Memorial Pier at SIO, located in La Jolla, California. Air samples and blanks were collected for our airborne dataset using The Coriolis Compact air sampler (CC) (Carvalho et al., 2008). Three sample types were used for analysis: two from the environment (Atmospheric and Seawater) and one blank from the instrument. Our seawater data came from the Southern

California Coastal Ocean Observing System (SCCOOS) (Terrill et al., 2006) data set, which was gathered biweekly from the Scripps Pier (lat:32.867, long: -117.257).

Atmospheric sample preparation:

The Coriolis Compact (CC) is a dry sampler with no liquid medium for sample collection. This sampling device employs polypropylene conical chambers to draw air into a sampling cone via a motorized inlet and then vortex the air into the chamber, collecting aerosolized particles on the cone's walls. Particulates, in our case microbes, are forced to the bottom of the cone, along the chamber's walls, where they accumulate. The sampler's design mitigates one of the primary challenges of airborne microbial studies, cellular abundance.

Airborne microbes have low biomass and cellular abundance. Therefore, our method required thorough cleaning and sterilization to prevent cross-contamination. To remove contaminants from the sampling cone, we cleaned the cones with UV sterilized and 0.2 μm filtered milli-Q water. The cones were then rinsed in a 10% hydrochloric acid solution to remove accumulated carbon deposits. This process was repeated twice, followed by a final rinse with Milli Q water. We further sterilized the CC sample cone in two phases. First, the sample component was sprayed with 70% ethanol solution and left to dry for two minutes inside a laminar flow PCR-grade hood. We then UV sterilized the ethanol-spritzed sampling apparatus for 15 minutes in a PCR hood equipped with a UV light. Concurrently, a sheet of aluminum foil was sterilized using ethanol and UV light. Our sampling cones were wrapped in this aluminum foil to maintain sterility in transport to the sampling site.

We affixed a retractable pole to the north end of Scripps Pier. We reduced the interference from foot traffic and pollutants detected on the pier by extending the sampler as far off the pier as feasible. Our instrument delayed collecting our samples for two minutes to reduce

contact with people and the terrestrial surface during the sampling period. Once the program began, it sampled the air for 6.5 hours at 50 L/min. The total volume of air collected was 19.5m^3. Blank samples were prepared the same, but the air sampler was deployed within a sterile PCR hood every seventh day. In total, five blanks and 30 atmospheric samples were collected. We collected the sample cone for each sample and took it to the laboratory for analysis. We recovered the biomass from the cone by rinsing the cone's walls with autoclaved milli-Q water. We collected 10mL of suspended material into a sterile 15mL falcon tube for each sample.

Seawater sample collection

Our seawater samples came from the SCCOOS program at The Scripps Ecological Observatory (SEO; Bowman et al., 2021). The SCCOOS program is a regional observing system that collects, integrates, and delivers coastal and ocean information via biological, chemical, and physical observations. At Scripps Pier Shore Station, microbial populations are analyzed biweekly using the sampling described by Wilson et al. (2021).

FCM sample collection:

We filtered 1 mL of homogenous unfiltered suspended material for our atmospheric samples through a 60 µm Supor membrane disc filter (Pall Corporation, Port Washington, NY, USA) into a 2 mL centrifuge tube. Each sample also received 10 µL of glutaraldehyde to biologically fix the sample in preparation for long-term storage at −80 ºC until processing with the flow cytometer.

For both Seawater and Airborne samples, we stained at the manufacturers recommended working concentration the 60 µm filtered samples solution of the nucleotide dye SYBR Green 1 (Molecular Probes, Inc., Eugene, OR, USA). We interrogated the cells with a blue (488 nm) and violet (405 nm) laser on a Guava easyCyte 11HT (Luminex, Austin, TX, USA). We performed a

complete cell-count quality control validation by spiking 10 µL of 1:10 diluted 1 µm 123 eCount beads (Polyscience Inc., Fishers, IN, USA) into each sample. We measured forward scatter, side scatters, and green emission (488/533 nm excitation/emission) at a flow rate of 14 µL min^{-1} for one minute. Following the protocol of Wilson et al. (2021) we identified grouping using a self-organizing map (SOM) from forward, side scatter and green emission from Syber Green stained nucleotides.

DNA sample collection:

While all seawater, airborne, and blank samples were processed using similar methods, the volume collected varied based on sample type. The total volume of the seawater samples was 1 L versus the 10 mL portions of processable material collected for the atmospheric and blank samples. All samples were analyzed for the bacterial and archaeal community structure using DNA sequencing and for microbial abundances with FCM.

We extracted DNA from our samples by filtering the remaining 8 mL airborne sample and ~1 L of seawater through a sterile 0.2 µm Supor membrane disc filter (Pall Corporation, Port Washington, NY, USA). We then placed the 0.2 µm filter into a bead-beating tube and stored the filter in the bead-beating tube at −80 ºC until extraction.

We assessed community structure through bacterial and arecheal16S rRNA gene analysis. DNA from lysed cells was extracted and purified in the KingFisherTM Flex Purification System and MagMax Microbiome Ultra Nucleic Acid Extraction kit (ThermoFisher Scientific, Waltham, Massachusetts, USA). The 0.2 µm filters were housed in the bead beating tubes and received 800 µL MAGMAX Microbiome Lysis Solution ™. The mixture was vortexed and then attached to a vortex plate for the bead beating lysis stage. Once fully lysed, we centrifuged the samples at 14G for 10 min. The supernatant contained unpurified DNA.

We processed approximately 400-500 µL of supernatant using the KingFisherTM Flex Purification System. We sent extracted and purified DNA to Argonne National Laboratory for amplicon library preparation and sequencing using the Illumina MiSeq platform with the universal primers 515F and 806R (Walters, 2016), a 2 x 151 bp library architecture.

DNA data processing:

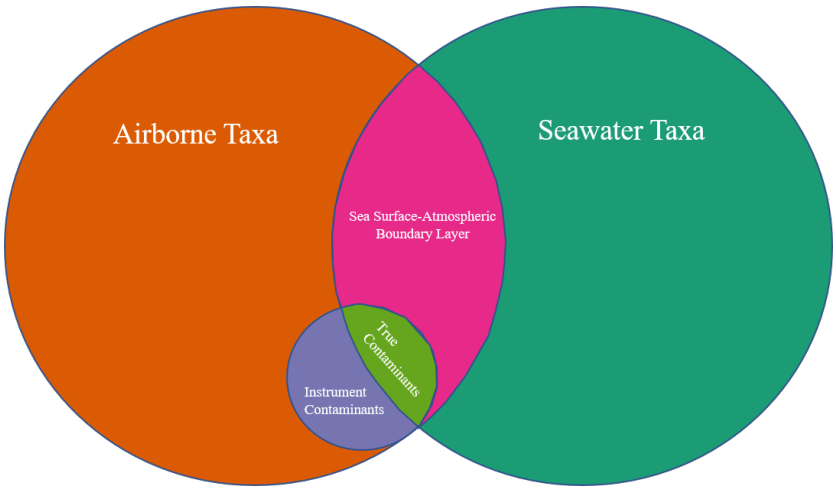

Figure 2: Airborne taxa are associated with atmospheric sampling, and marine taxa are associated with Seawater samples. The overlapping community is the Sea Surface-Atmospheric Boundary Layer. Instrument contaminants were removed from taxa pool.

After receiving our reads from Argonne National Laboratory, we first used dada2 to filter, denoise, and merge the readings before analyzing them with PAthway PRediction by phylogenetIC plAcement (paprica) v0.7.1(Bowman and Ducklow, 2015). As part of the phylogenetic placement process, paprica utilizes Gappa (Czech, Barbera et al. 2020), EPA-ng (Barbera, Kozlov, et al. 2019), Infernal (Nawrocki and Eddy 2013), as well as RefSeq (Haft, Dicuccio, et al. 2018) within the pipeline to place reads on a reference tree constructed from all completed genomes in GenBank. The reference tree assigns all unique reads to internal or terminal branches. Finally, we deleted reads linked to metazoan mitochondria or chloroplasts and sequenced reads that occurred only once.

Blank Contamination Investigation

In our investigation, we collected five blank samples total, one sample per week, and used them to identify contamination brought into our experiment from the CC air sampler. Four of the 79 taxa found in all blank samples were not actual contaminants. We identified ASVs with significantly higher mean relative abundance in seawater samples than expected due to chance and contamination alone ($p = 0.01$). We identified these as probable environmental taxa, and they were returned to the taxonomic pool.

Subsetting the SSABL:

As part of our research, we determined which taxa are common to seawater and airborne samples in the SSABL layer. We subset our data into three categories: airborne seawater and blank samples. We identified the SSABL as the subset of common taxa found in seawater and airborne samples, with all contaminants removed from the blanks. After removing the blank taxa from our sample pool, we identified 3761 taxa across all samples.

HYSPLIT:

To source the airmasses that brought our samples to the pier, we used an airmass trajectory model from NOAA (Stein et al., 2015). This method was adapted from Xia et al. (2015), which described how to integrate a series of trajectories over time. Our analysis was conducted over a 35-day window. The NOAA Air Resources Laboratory's Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) transport and dispersion model (available online at http://www.arl.noaa.gov/ready/hysplit4.html) constructed airmass backward trajectories (BTS) for all 35 days of the airborne sampling collection window.

The HYSPLIT model collected airmass data from a preselected window of 13:00UTC-21:00UTC at our endpoint of: lat:32.867, long: -117.257 (North end of Scripps Pier). The model

performed 256 trajectories, 8 per day from May 10th to June 13th (32 days), running a new air mass trajectory every three hours and retracing the trajectory for 72 hours. The model measured each trajectory at fifty meters above the mean sea level 0.25x0.25degree resolution with the NAM-12km data archive from 13/6/21. HYSPLIT software generates a frequency distribution of the air masses and outputs a heatmap of air mass distribution. Each frequency distribution equals 100 * (the number of endpoints per grid square)/(the number of trajectories.)

Statistical analysis:

Differential Abundance: To test differential abundance, we looked at the taxa within the SSABL and performed a series of Wilcox tests comparing the mean relative abundance of all taxa found in the SSABL between all samples. We used a Bonferroni-adjusted p-value to select the most significantly differentially abundant taxa within the SSABL.

Diversity Metrics:

We measured alpha diversity through Shannon's Diversity metric, Inverse Simpson Diversity, Species Richness, and Species evenness. We calculated our diversity metrics and species richness using the vegan package in R. Species evenness is measured by the Shannon diversity metric divided by the log of our species richness. We calculated the same metrics for the SSABL. For each data set, we compared Atmospheric samples to Seawater samples.

PCoA

We created a Bray-Curtis dissimilarity matrix of our atmospheric and seawater samples and measured the distance between samples using the vegan package. The cmdscale function from vegan evaluates a pairwise assessment of the Bray-Curtis dissimilarity between samples within ecological groupings and attempts to scale the relationship into low dimensional space. Each axis represents the explanatory power of that relationship.

Correlation of Taxonomic Groupings to PCo1.

Determination of which taxa have the most significant effect on PCo1, the primary axis of our ordination, we used a spearman correlation to correlate PCo1 to the relative abundance of each taxon in our taxonomic pool. We selected the 10 most significantly correlated taxa (p-value <0.05) with a relative abundance greater than 1% of all taxa and a rho greater than an absolute value of 0.5. We vectorized the effect of these taxa onto our PCoA.

Similarly, we performed the same analysis of our environmental variables and correlated the effect of these variables on PCo1. We removed variables that were autocorrelated with each other (ex. Minimum solar radiation and maximum solar radiation) and selected 12 possible variables that have some effect on PCo1. We used these variables in our multiple regression analysis.

Multiple Regression Analysis:

To assess how the airborne community structure is affected by physical properties, we utilized a multiple regression on the normalized PCo1 values as our response variable and the following variables as predictor variables:

Average wind speed (kn), average Wind Gust (kn), minimum barometric pressure (Mb), average verified sea height (ft), average water temperature (deg C), average offshore & longshore wind direction, average solar energy (W/m2), average UV index, average humidity, average dew point, and average air temperature.

We then systematically removed variables from our model to generate the best possible Akaike Information Criterion (AIC) score. Our final model's parameters were:

Average verified sea height (ft), average water temperature (deg C), average offshore & longshore wind direction, average solar energy (W/m2), average UV index, average dew point, and average air temperature. To assess if our model was overfitting, we used a bootstrapping method to run 1000 iterations of the model, withholding 3 random parameters each time and assessing the level of fit. Lastly, we took the parameters of our model and vectorized the effect of the variables onto our PCoA.

Results:

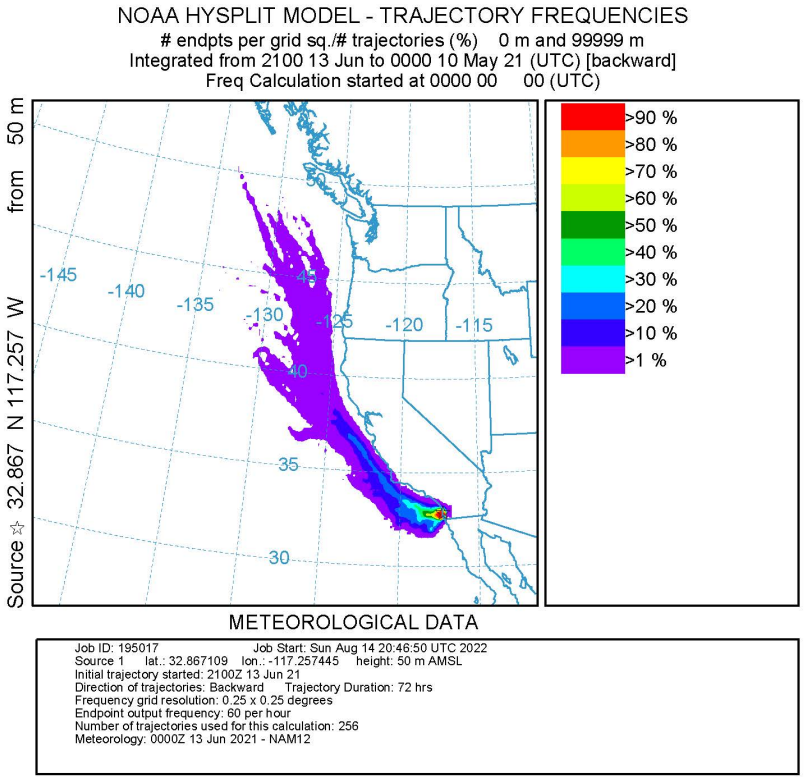

Figure 3: *HYSPLIT Frequency Analysis.*
Figure Depicts Backwards Trajectory Analysis for data collected from May 10th 2021 to June 11th 2021. Each Trajectory starts at the Scripps Pier and is traced backwards for 72 hours.

Hysplit:

Air mass backward trajectories analysis reveals that all samples collected during the investigation are primarily from oceanic air masses (Fig 3) Air masses followed the United States and Canadian continental margins, with ninety percent of trajectories never passing over land.

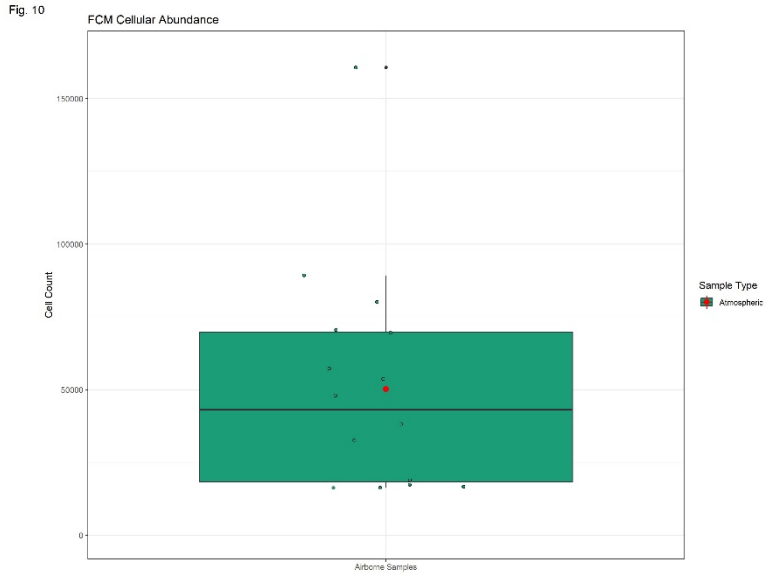

Figure 4: Flow Cytometry Counts for Atmospheric Samples.

Community Composition

FCM:

The airborne component's overall abundance is considerably lower than that of marine or terrestrial environments (Burrows et al., 2009). Our measurement of cellular abundance is shown in Figure 4. The Bertin CC draws in air at a rate of 50L per min. We sampled for 6.5 hours for 19,500L of air or 19.5 cubic meters.

With an average cell count of 50236 cells per sample, we estimate an average of 2576 cells/m^3 (cells per meter cubed).

Abundance

In the heatmap of all taxa collected, we identified the top fifty most abundant taxa collected in our investigation and plotted their relative abundance per each sample. The samples labeled “sccoos” are our seawater samples, and the samples labeled “blamm” are our airborne samples. The numeric tag following the sample type designation is the date of sampling. Samples

with a higher abundance of a particular taxon are shown in dark green, and the samples with lower relative abundance are shown in yellow.

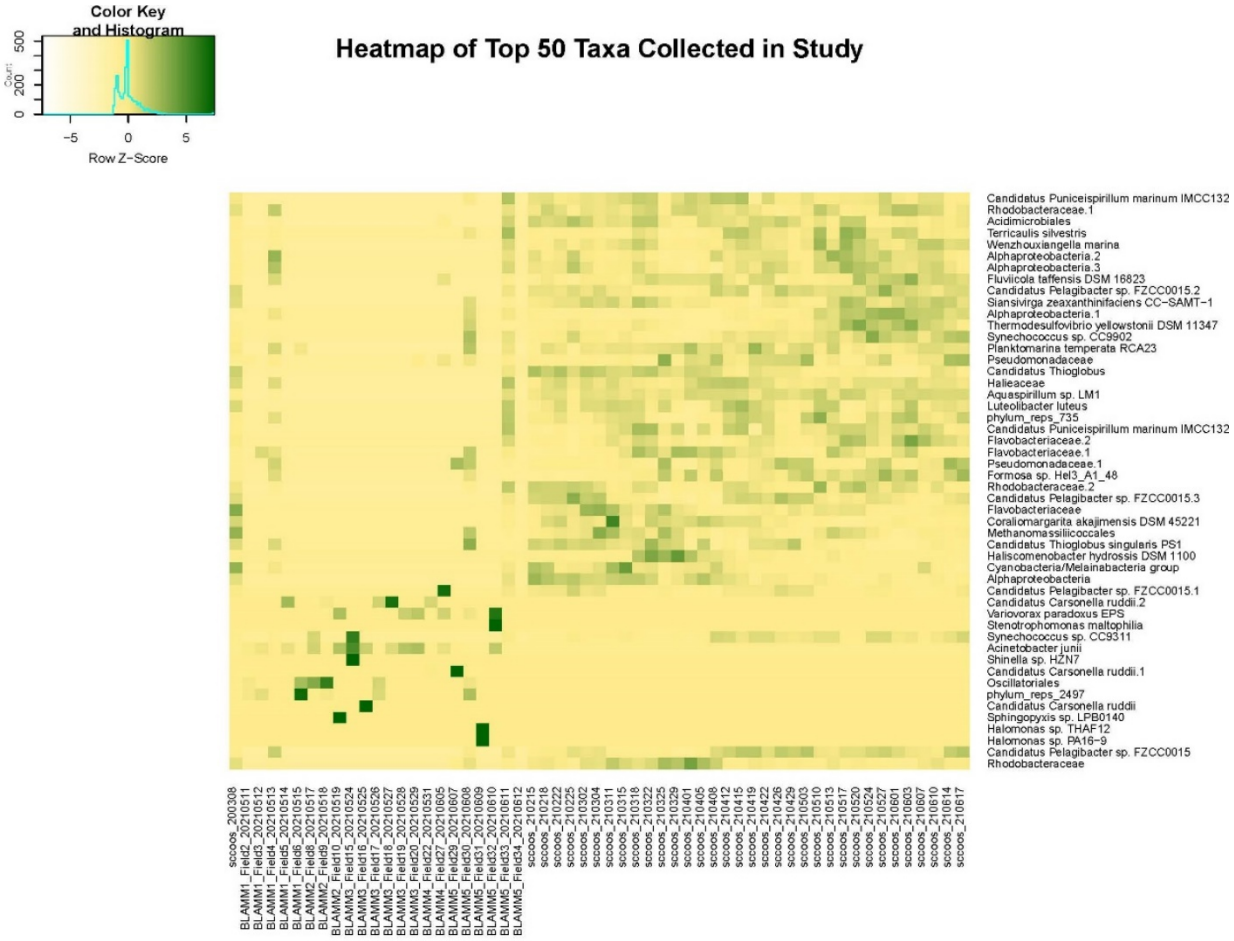

Figure 5: Heat Map of Relative Abundance of the Top 50 most abundant Taxa collected.
Note: Atmospheric Samples are Denoted "BLAMM" and Seawater Samples are denoted "sccoos."

SSAB Abundance

We define the composition of the Sea Surface-Atmospheric Boundary Layer community as the taxa common to both the sea surface and the atmosphere. The taxa do not have to be in every sample, but they have to be common to both environmental groups. We again use a heatmap to visualize the top fifty taxa in this ecological space.

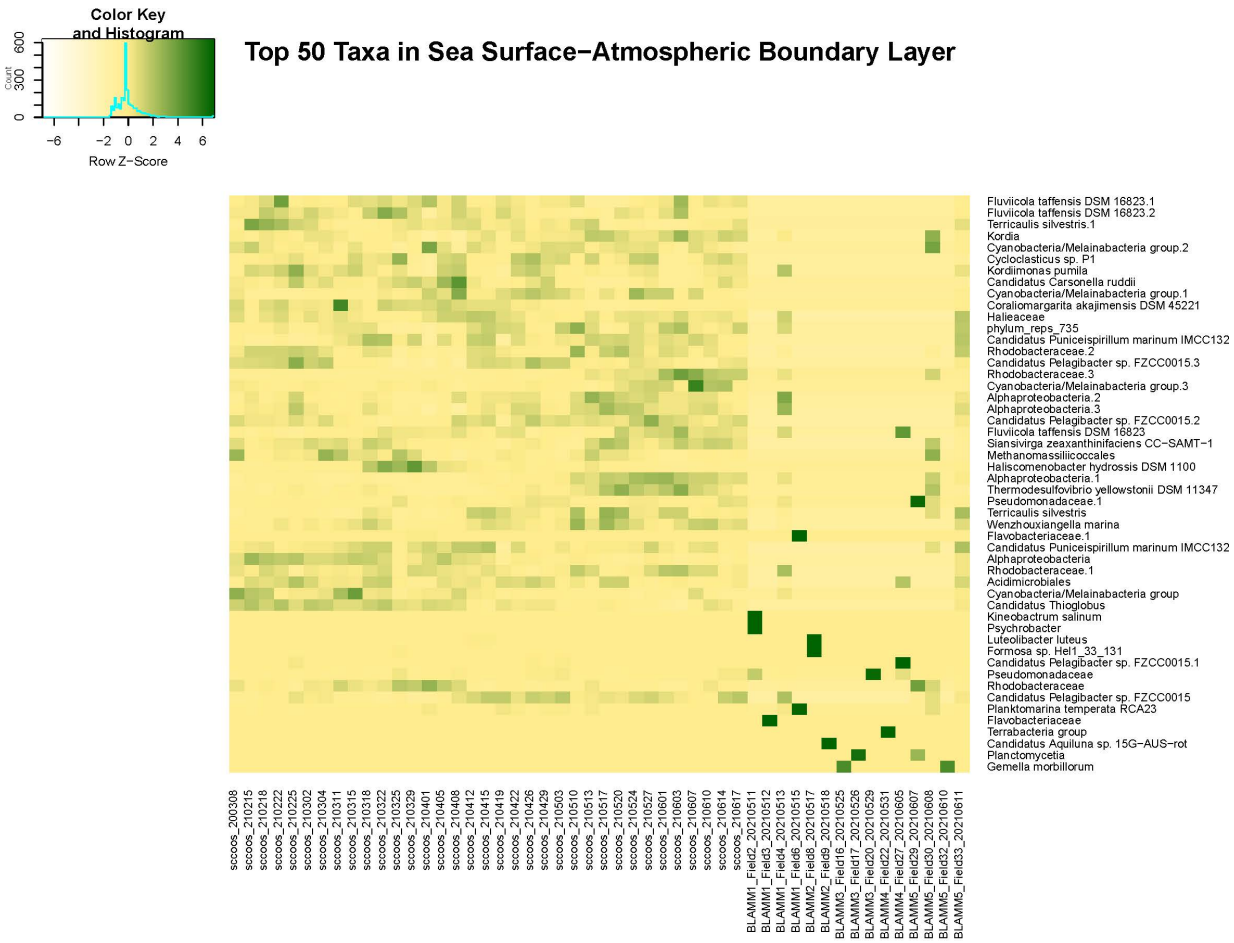

Figure 6: Relative Abundance of Top 50 Taxa found in the SSABL.

Our diversity measurements displayed significant differences between sample type, with the seawater samples being considerably more diverse, and offering a more robust and rich community than that of the SSABL. Within the SSABL, the diversity indexes of the airborne samples are markedly lower than in our seawater samples. We observe this same trend in our species' evenness and richness_diagrams.

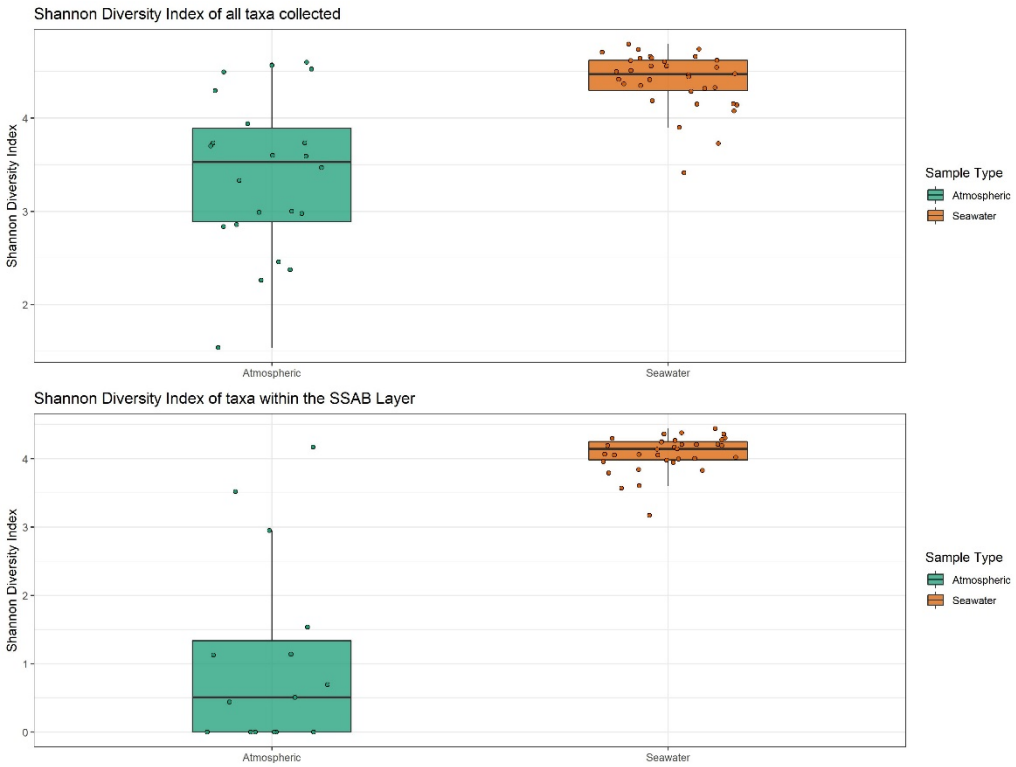

Figure 7: Shannon Diversity Index.
Comparison of two sample types. Top chart displays diversity for all taxa collected, and the lower chart depicts the diversity of the SSABL.

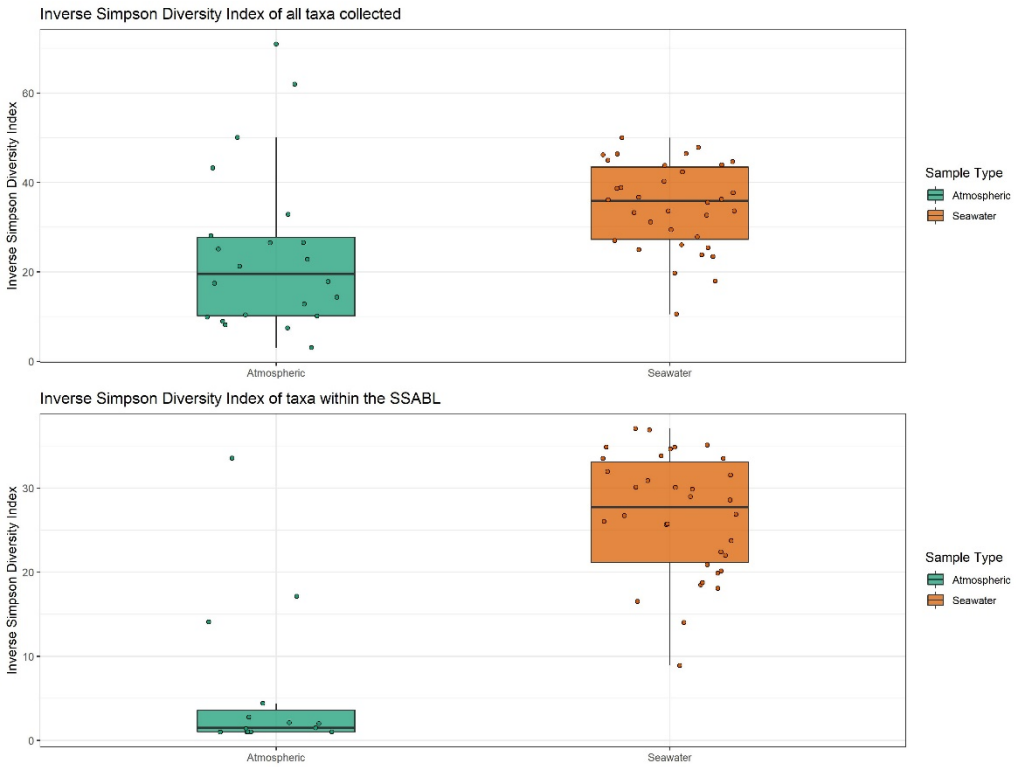

Figure 8: Inverse Simpsons Diversity Metric*:*
Top chart displays diversity for all taxa collected, and the lower chart depicts the diversity of the SSABL.

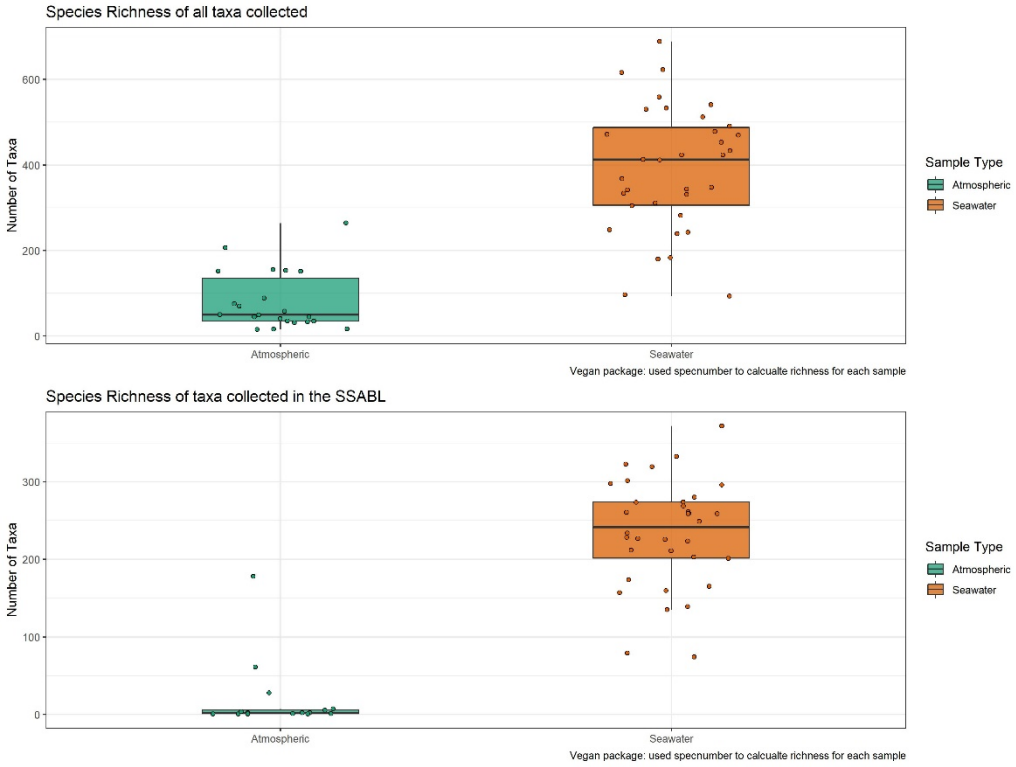

Figure 9*:* Species Richness Comparison.

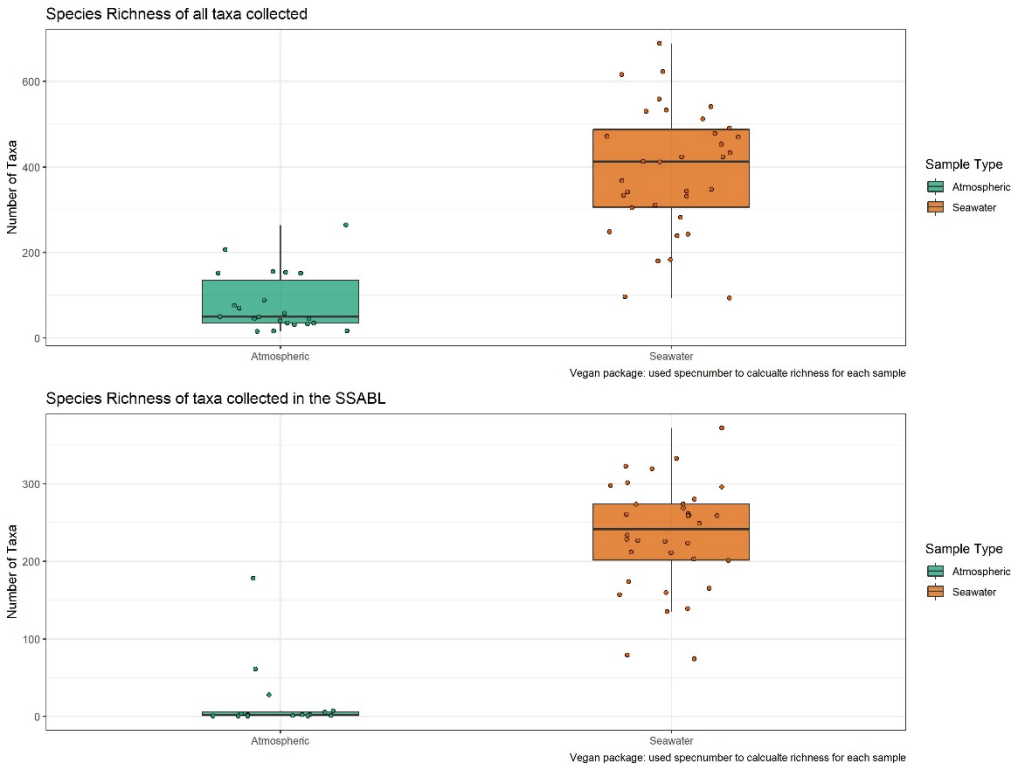

Figure 10: Species Evenness Comparison.

Table 1: Alpha Diversity Results

Diversity Metrics Results								
Average results for each diversity metric and standard deviation								
Sample Type	shan avg	inv simp avg	rich avg	even avg	shan std	inv simp std	rich std	even std
Atmospheric	3.403	24.085	80.909	0.834	0.831	18.109	68.562	0.098
Seawater	4.401	34.631	391.706	0.750	0.302	9.646	146.254	0.025
Atmospheric ssabl	1.073	5.660	19.667	0.485	1.392	9.193	46.672	0.423
Seawater ssabl	4.077	26.792	233.471	0.756	0.263	7.161	69.245	0.024

Table 2:Kruskal-Wallis Rank Sum Test Results

Kruskal-Wallis rank sum test				
Comparison of Diversity metrics (P-Values)				
Sample Type	Shannon Diversity	Inverse Simpson Diversity	Species Richness	Species Evenness
Airborne vs. Airborne SSABL	4.89e-05***	6.365e-05***	6.333e-05***	0.01314*
Seawater vs. Seawater SSABL	3.148e-06***	0.00065***	1.773e-06***	0.162
Signif. codes:	0 ‘***’	0.001 ‘**’	0.01 ‘*’	0.05 ‘.’

The Kruskal-Wallis rank sum test indicates that all sample types had significantly different diversity metrics when subset to the SSABL, except for Species evenness in the Seawater samples where no significant difference was observed.

In Figure 11 we show the taxa that make up the SSABL community composition and compare the relative abundance of those taxa across sample types. The taxonomic groups that make up less than 1.25% are pooled into one category labeled “Other”

diversity metrics

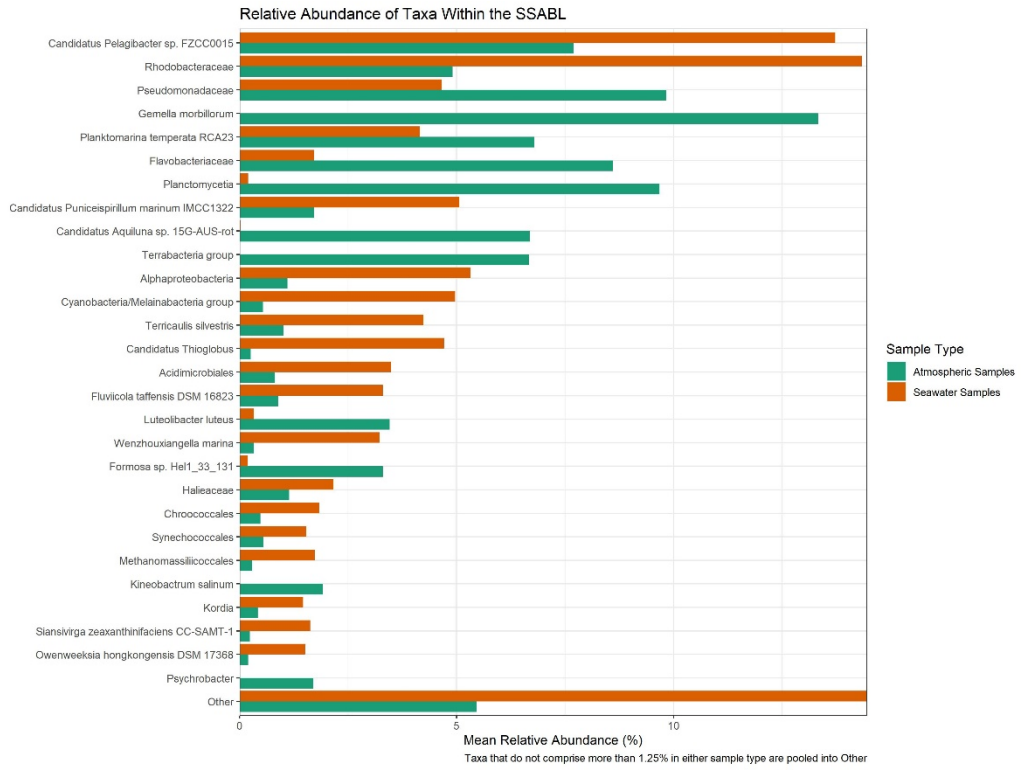

Figure 11: Relative Abundance of Taxa within the SSABL.

Highlighting community Distinctions:

Community Composition:

The evaluation of differential abundance of the SSABL taxa (Figure 12). Each comparison comes from a paired Wilcox test measuring each taxon's abundance across our environmental sample types. The taxa shown in the figure have a p-value of less than 0.05. If a taxon is differentially abundant, its mean relative abundance must be statistically different between the two ecosystems. Sixty-seven taxa in the SSABL had a significantly different mean relative abundance (p-value >0.05) between data sets. Figure (3) shows the thirty-four taxa with a highly significant p-value of less than 0.001.

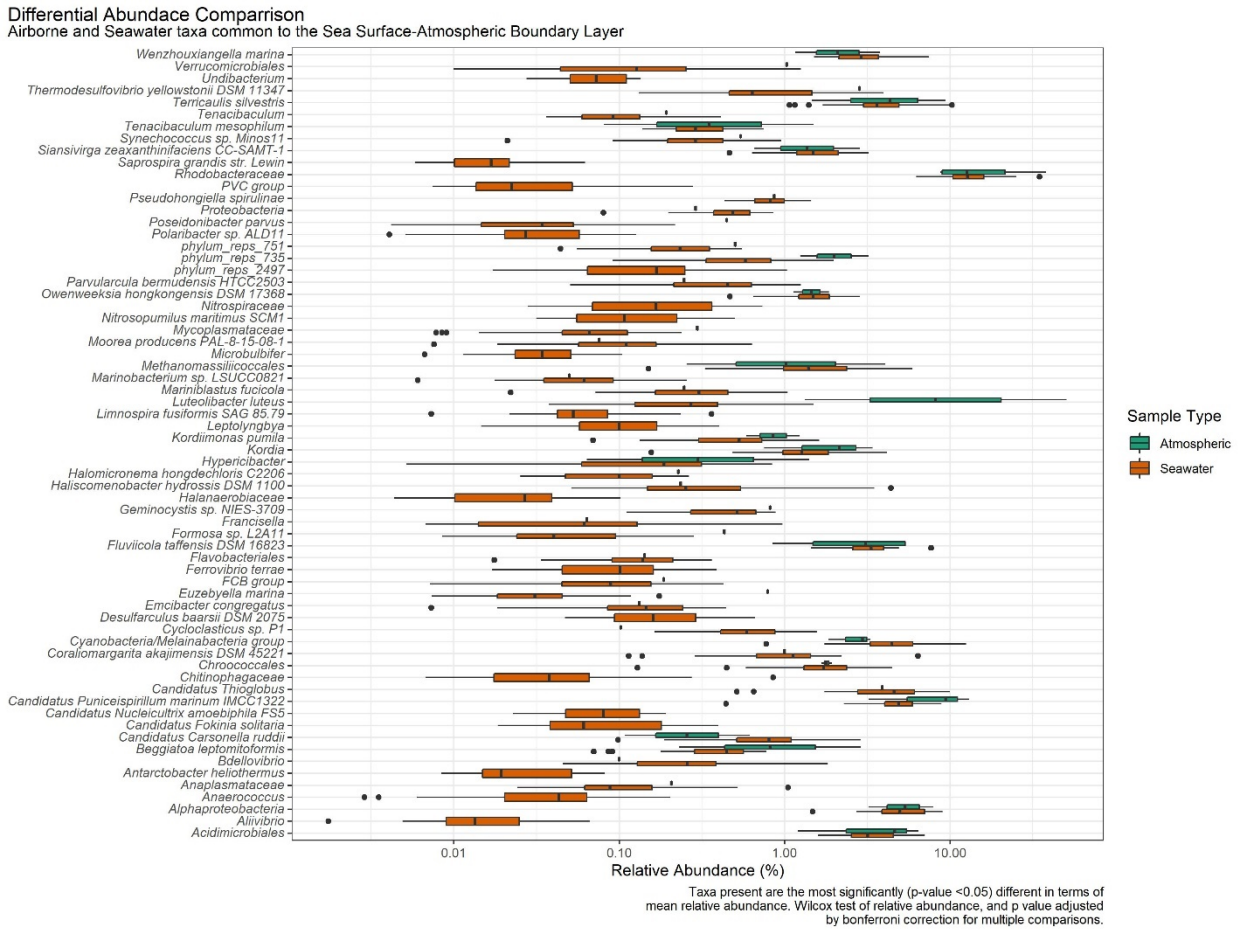

Figure 12: Differential Abundance of Taxa within the SSABL.
Differential abundance calculated using a paired Wilcox test, and p-values were Bonferroni Adjusted for multiple comparison.

Ordination and distinction of community sets:

The ordination is the distance between samples evaluated by a pairwise assessment of the Bray-Curtis dissimilarity between samples within ecological groupings. Figure # (PCoA) is a Principal Coordinates Analysis (PCoA) ordination of the relative abundance of the marine taxa (collected from our seawater samples) and our airborne taxa (collected from our atmospheric samples) and highlights the distinction between the two ecological groupings

Table 3: ADONIS Betadispersion Results

	Df	SumOfSqs	R2	F	Pr(>F)
Sample Type	1	4.9744	0.28424	21.444	9.999e-05***
Residual	54	12.5264	0.71576		
Total	55	17.5008	1.00		
Signif. codes:	0 ‘***’	0.001 ‘**’	0.01 ‘*’	0.05 ‘.’	0.1 ‘ ’
	Number of permutations: 1000				

The two communities are markedly distinct, with a few exceptions seen in a few airborne samples. These samples have a higher degree of similarity in terms of community composition.

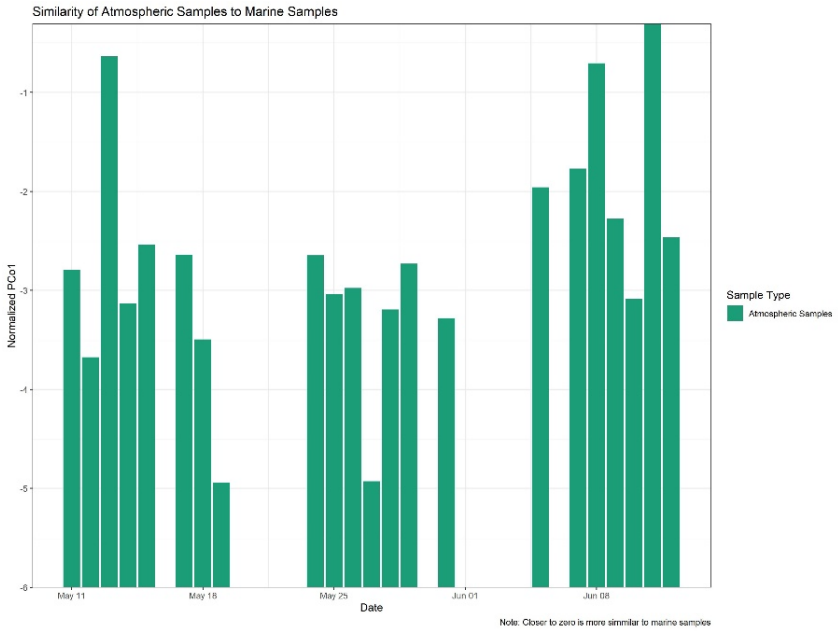

Figure 13: Log Normalized PCo1 values of airborne samples over time.

We looked for a metric to quantify how similar our airborne samples were to the marine samples. Figure 13 shows a log normalized value from PCo1 for each airborne sample collected during our sampling period. These values are an estimate of how similar the airborne taxa are to

the marine samples. The closer to zero implies a greater degree of similarity. Days of highest degree of similarity were: May 13th 2021, June 8th 2021, and June 11th 2021.

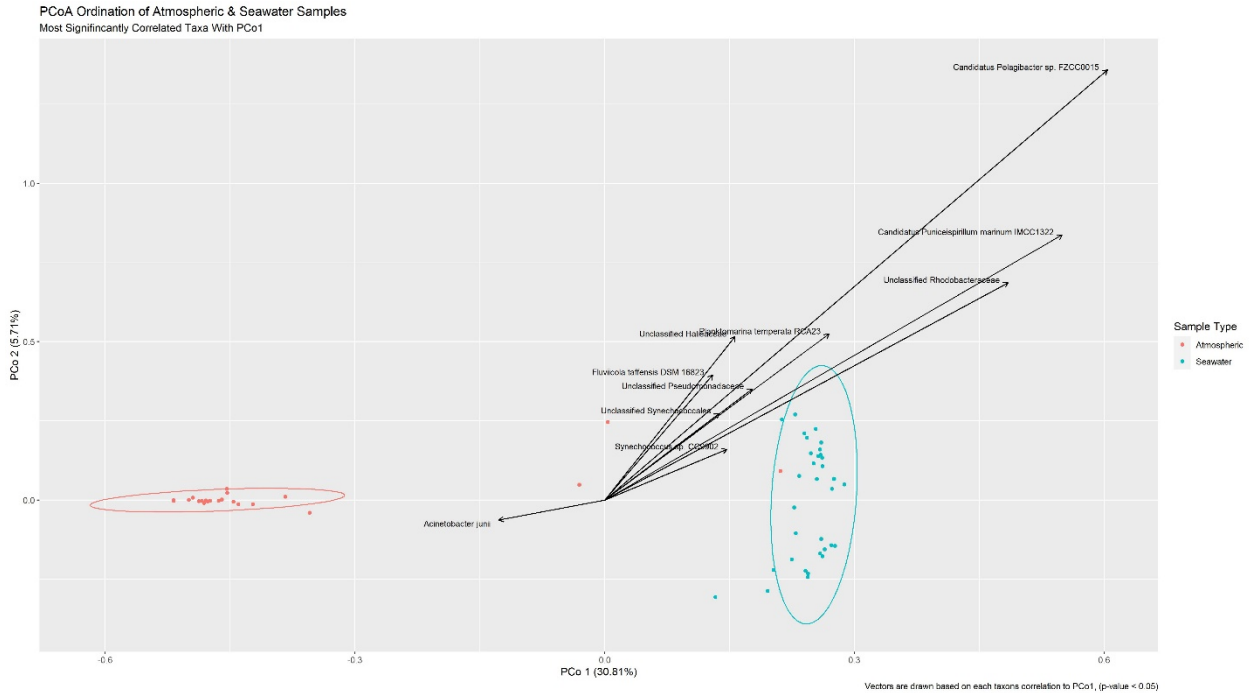

Figure 14: PCoA Ordination with Ten most correlated taxa to PCo1

Correlation of taxonomic abundance to PCo1:

In our taxonomic correlation analysis, we identified ten taxa with significant effect of correlation on PCo1, these were the taxa seen above the redline on Figure 15.

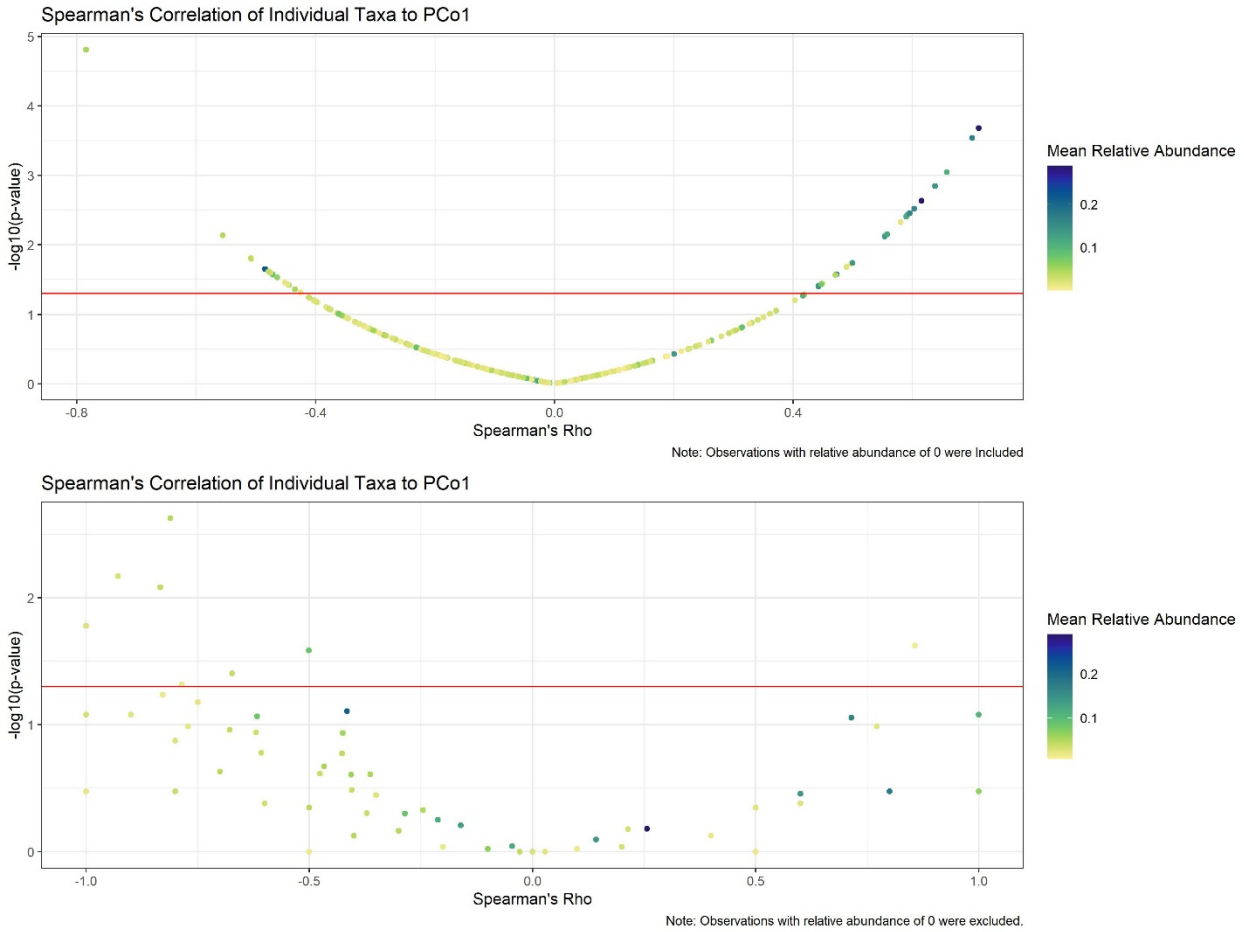

Figure 15: *Spearman's Correlation of taxa to PCo1.*

The taxa Identified were:

Table 4: Significantly Correlated Taxa

Acinetobacter junii	Planktomarina temperata RCA23
Candidatus Pelagibacter sp. FZCC0015	Unclassified Pseudomonadaceae
Candidatus Puniceispirillum marinum IMCC1322	Unclassified Rhodobacteraceae
Fluviicola taffensis DSM 16823	Unclassified Synechococcales
Unclassified Halieaceae	Synechococcus sp. CC9902

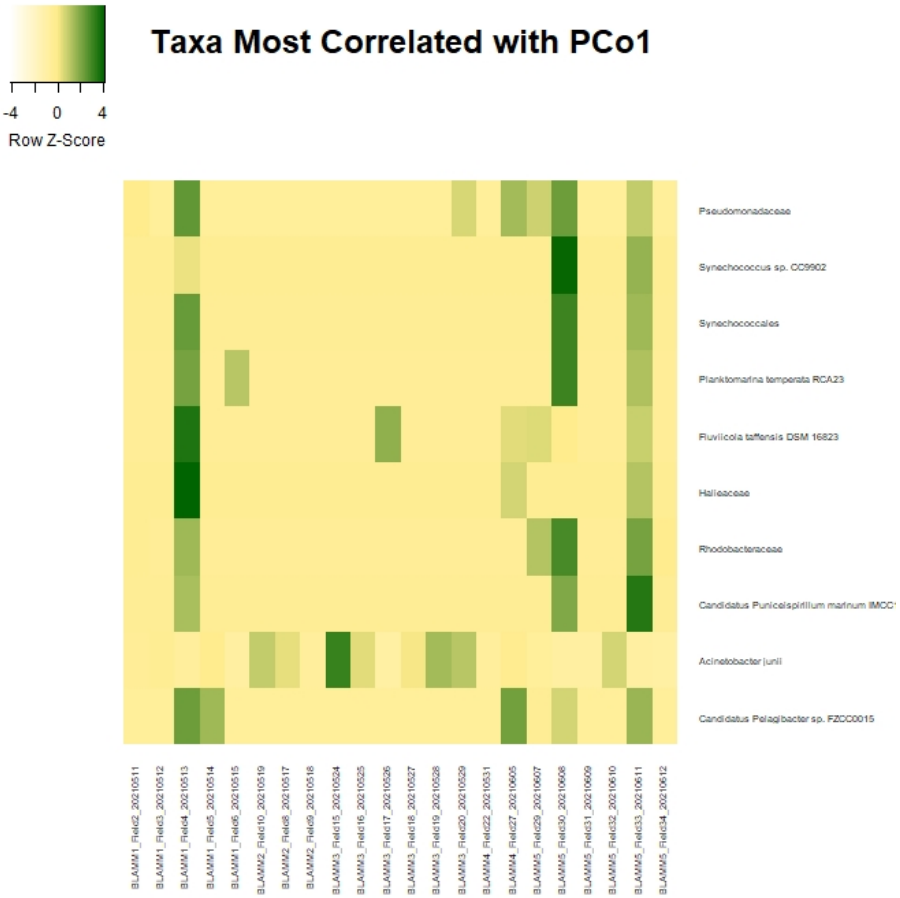

Figure 16: Relative Abundance of Taxa correlated with PCo1 Within Atmospheric Samples

The taxa were identified and vectorized, then overlaid onto our PCoA. The direction and magnitude of the vector indicated the effect the taxa have in scaling the ordination (**fig 14)**.

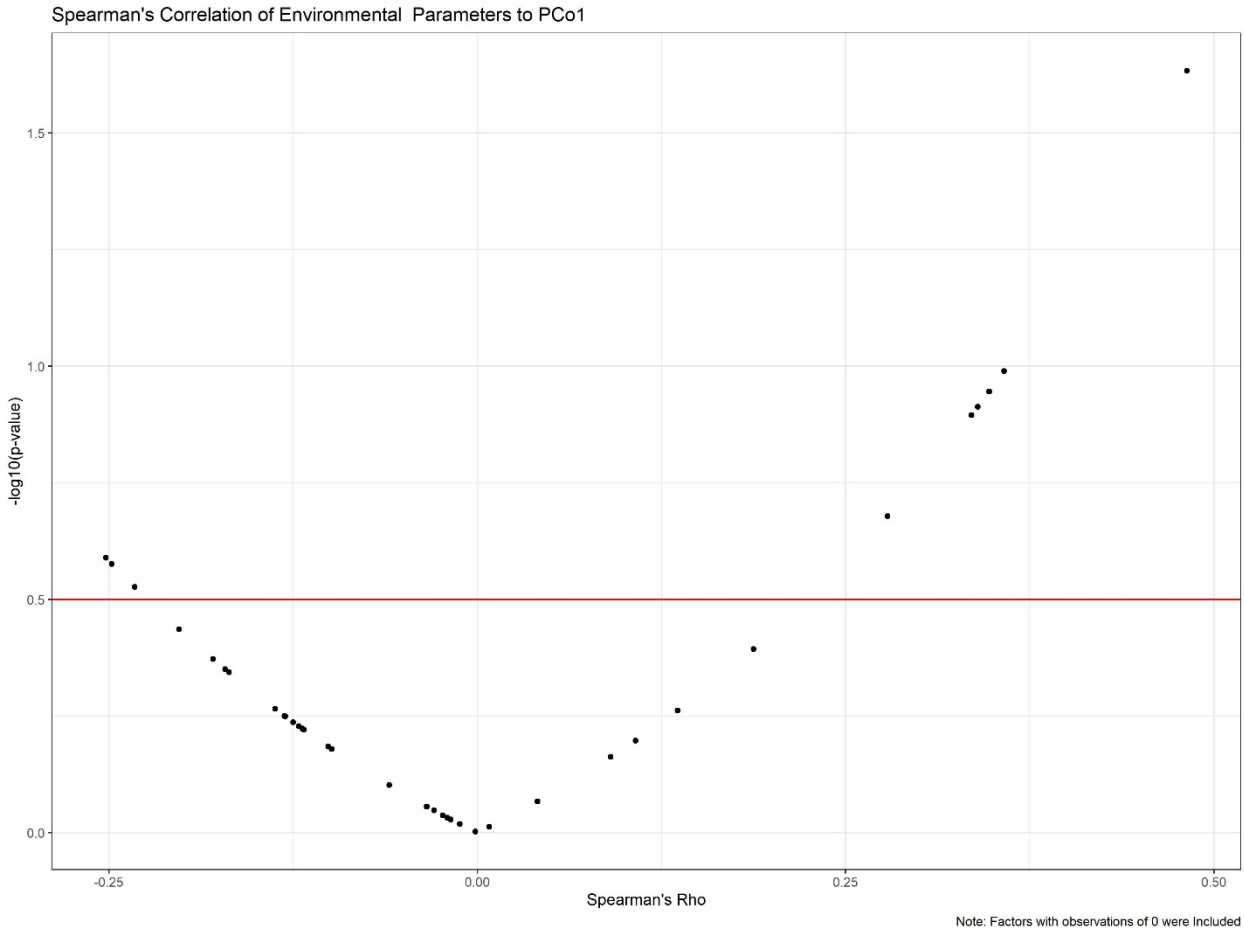

Figure 17: Spearman's Correlation of Environmental Variables to PCo1

Physical drivers of differentiation

Figure 18: PCoA ordination of samples with correlated environmental factors.

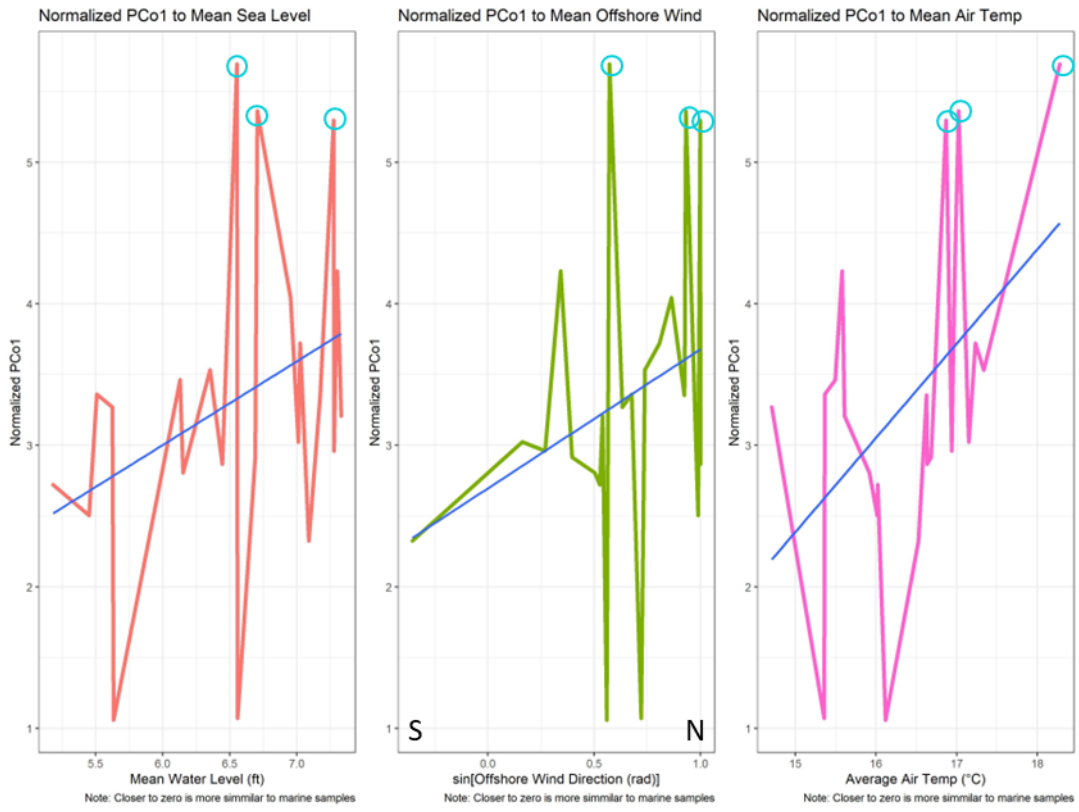

Figure 19:Airborne samples that are more marine like are found on days with: Higher Mean water level. Northward wind direction, and Warmer Avg Air Temperature.

Linear regression:

Our model displayed an adjusted R^2 of 0.766, with several of our predictors having a significant effect on our response variable (p-value < 0.05) on PCo1.

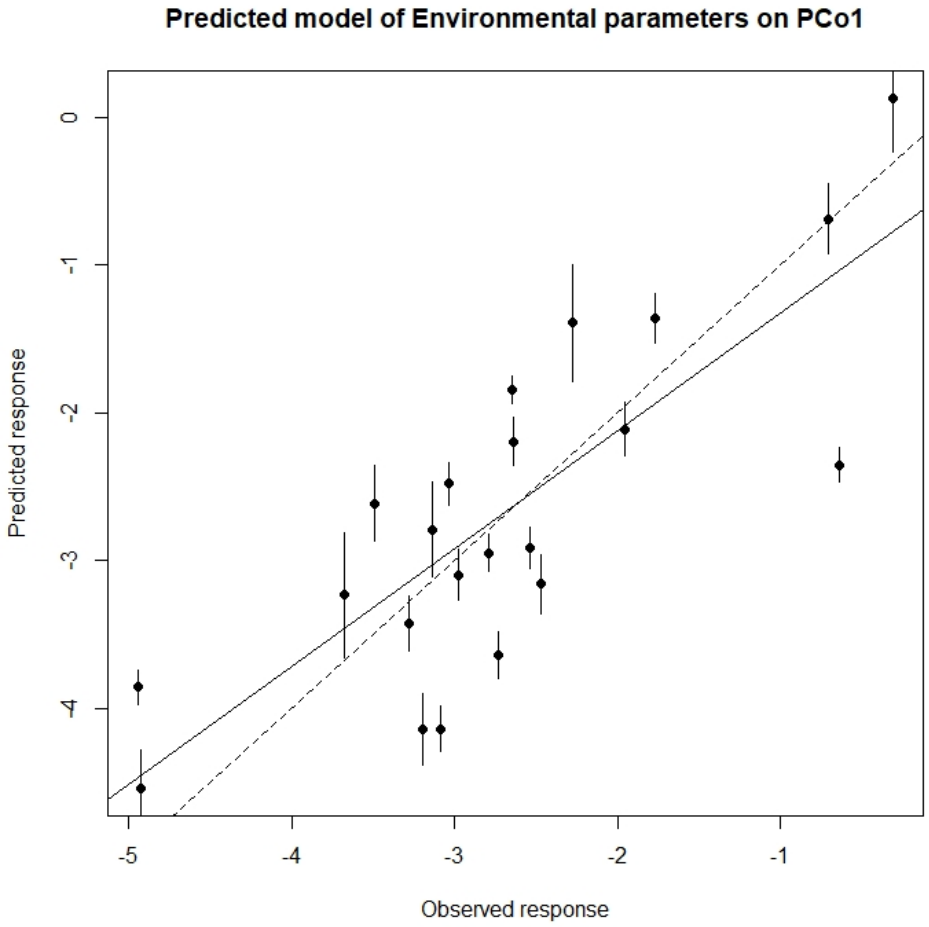

Figure 20: Predicted Vs Observed Model. Solid line is slope of regression, and dashed line is 1:1 fit.

Table 5: Physical Parameters that effect PCo1

Multiple Regression Analyis

Formula:
Lm(normalize_marine ~ avg_Verified.ft + avg_Water.Temp.degC + avg_offshore_wind + avg_longshore_wind + avg_solar_nrg + avg_uvindex + avg_dew + avg_Air.Temp.degC)

Term	Estimate	std.error	Statistic	p.value
(Intercept)	-41.626	6.166	-6.751	1.36e-05
Avg Verified Sea Height	**0.489**	**0.219**	**2.236**	**0.044**
Avg Water Temp (C)	**0.541**	**0.188**	**2.880**	**0.013**
Avg Offshore Wind Direction	**2.416**	**0.716**	**3.374**	**0.005**
Avg Longshore Wind Direction	**-1.158**	**0.509**	**-2.275**	**0.040**
Avg Solar Energy	9.824	4.620	2.126	0.053
Avg UV Index	**-4.555**	**1.651**	**-2.758**	**0.016**
Avg Dew Point	-0.304	0.145	-2.097	0.056
Avg Air Temp (C)	**1.929**	**0.280**	**6.890**	**1.10e-05**

We vectorized the parameters of our linear regression to fit our PCoA ordination, to better understand how the effects of these variables can affect the similarity to a marine microbial community (Figure 18).

Discussion

We present the reader with an opportunity to delve into a microbiome that we have learned is complex, dynamic, and fickle.

The Sea Surface-Atmospheric Boundary layer is akin to looking at the sea surface from a desert. Directly in front of us is a vast oasis teeming with life, but in the air directly above the sea surface, only a few well-adapted species can survive in a harsh environment.

The understanding of the airborne microbiome is that there is comparatively low biomass in the air than its terrestrial and aquatic counterparts (Burrows et al., 2009). The atmospheric physical forcings predict a low microbial abundance (Dueker et al., 2011). Wind and surrounding ecosystems are excellent predictors of microbial communities (Tignat-Perrier et al., 2019). Most studies have employed large aerosol samplers to overcome these challenges (Dommergue et al., 2019). In our research, we had expected to encounter low biomass (Archer et al., 2020).

In our study, we wanted to employ a relatively affordable and portable piece of equipment, The Bertin Coriolis Compact (Carvalho et al., 2008). The device is roughly the size of an average coffee machine and has a rechargeable battery, making it a valuable tool for situations where bulky and heavy samplers are onerous to deploy. With this tool and a unique position of working within SIO, we developed an experiment to build an ecological profile of the SSABL using the CC.

Our research focuses on assessing how similar the microbial community of the atmospheric samples to the seawater samples are. We sought to characterize the microbial community of the SSABL, identify significant taxa present within the ecosystem, and describe the physical properties that can affect the community structure. We hypothesized that given the

proximity to the sea surface, the microbial community of the SSABL would be driven primarily by marine taxa.

As noted, microbial abundance was our most challenging obstacle to overcome. We managed to develop a plan that would mitigate abundance issues by forming a month-long sampling effort. With daily six-and-a-half-hour sampling periods, we collected samples for DNA and flowed cytometry and turned the equipment around for the next sampling effort in a day. Thus, we felt this design would capture a clear snapshot of the microbiome.

We formed an ecological data set from our processed data. The data set included DNA sequence reads and flow cytometry counts. Due to the low microbial biomass, not all of our samples could develop genomic libraries, nor could every sample be adequately counted. We cleaned our genomic data using the deseq package(Love et al., 2022) and, in turn, used the paprica pipeline (Bowman & Ducklow, 2015) to identify Amplicon sequence variants (ASVs) and process the microbial abundance data.

Previous research in the field has shown that tracking the trajectories of air masses yields sourcing information on microbial populations (Xia et al., 2015). Our study performed a frequency analysis of the air mass trajectories. We show that oceanic air masses flowed along North America's coast in May and June of 2021, arriving at our sampling site on the Scripps Pier. Therefore, we expected to see some degree of the marine microbial community within our airborne samples collected. In our analysis, we identified and described the community structure of the SSABL, which we use as evidence in testing if the marine community interacts with the SSABL and to what degree the taxa exhibit a significant effect.

Diversity

Taxonomic analysis of collected samples identified similarities between the airborne and marine samples, highlighting the connections between the two communities. Interestingly, despite similar phyla between the two environments sampled, there was variation in the specific taxa isolated from each environment, as seen in the heatmaps in figures 5 & 6. Also, The keen-eyed may discern that in the heatmaps (particularly that of the SSABL), very few taxa are represented consistently in the airborne samples (denoted BLAMM). The predominant pattern appears to exhibit only a few taxa within each sample, but those present are relatively abundant (darker green).

Verified by our diversity analysis, we see that the airborne portion of the SSABL's quantitative community metrics is significantly lower than that of their total airborne community(p-value). In contrast, the sea surface diversity metrics remain consistent when subset to the SSABL(p-value).

Our differential abundance testing of the SSABL also highlighted distinctions in the taxonomic make-up of the microbiome.

If a taxon is differentially abundant, its mean relative abundance is statistically different between the two ecosystems. Each taxon shown has a p-value less than 0.05. As shown in figure 14, we see many more differentially abundant taxa marine in origin. This phenomenon is likely due to either low reads or absence in many atmospheric samples.

This further drives the point that the marine component is the most influential piece of the SSABL.

The typical abundance of airborne bacteria above coastal waters averages 7.6×10^4 cells M^-3 [(Harrison et al., 2005)]). We measured a much lower average concentration at only 2576 cells per m^3, as seen in figure 6. Other studies have reported abundances lower than (number).

In our case, variability of the environment is a likely contributor; it is more likely due to the CC as the air sampler is designed to be portable and easily deployable but may be limited in its capabilities to collect concentrated biomass. The analysis of the communities' diversity, absolute abundance, and community make-up suggests that bacterial populations in each environmental type are distinct and that the airborne community is unique compared to the marine environment.

With such a different composition, we needed to address the taxonomic make-up of the two sample sets and compare the differences.

PCoA is a flexible method of handling complex ecological data. PCoA seeks to accurately represent the pairwise dissimilarity between objects in a low-dimensional space by condensing ecological abundance data along 2 to 3 axes of explanatory power. The indirect gradient analysis method, known as a technique for metric multidimensional scaling, PCoA creates an ordination based on a dissimilarity matrix. (*Principal Coordinates Analysis - GUSTA ME*, n.d.) (joshuaebner, 2018).

This ordination highlights one of the principal findings of our analysis: the two communities are distinct. Though some Airborne samples are more closely associated with the marine samples in ordinal space, the vast majority of airborne samples appear to have no similarity with the sea surface community. Our beta dispersion analysis shows that the two sample sets differed significantly (p-value).

The PCoA most accurately describes the community structure of the SSABL. The vast majority of atmospheric samples collected cluster together, with some exceptions closer to the seawater samples.

The question remains, what taxa and what environmental properties are driving this distinction? To address this issue, we correlated the relative abundance of each taxon to the primary axis of ordination, PCo1. Spearman's Rho estimates indicate the direction of correlation (i.e., positive and negative), while the y-axis shows significance. We identified the taxa with the highest correlation to PCo1 and a significance of p-value <0.05. We vectorized these values and displayed them on our ordination. The vectors display the magnitude and direction that these taxa affect our communities' ordination (Oksanen et al., 2020)—the taxa with the most significant effect on PCo1 heavily favor the marine direction toward the seawater samples.

Samples with a higher abundance of these taxa will be more similar to the marine environment within this ordination. We show this in Fig20, where the samples with the highest abundance of the taxa most associated with PCo1 are more abundant in the three atmospheric samples closest to the marine samples in the PCoA ordination.

We also correlated the environmental variables to the PCo1 to determine which physical properties influence the community distinction. We needed to develop a multiple regression analysis alongside the correlated parameters. Our Spearman's correlation analysis highlighted ten variables that affected the ordination of PCo1. We then used those ten factors in a multiple regression analysis as our explanatory variables to predict PCo1. After systematically removing variables from our model, eight variables remained that significantly affected the ordination of PCo1. When ordinated on the PCoA, we see that 6 of the eight factors orient themselves toward the marine samples. This ordination indicates that these are primary drivers in what makes airborne samples more marine-like.

In our analysis, we collected 35 atmospheric samples and attempted to build genomic libraries of the microbes within the airborne component of the SSABL. Though we used 21st-century DNA purification systems and next-gen sequencing techniques to develop these libraries, we could not collect genetic data on all 35 samples. We believe this issue arises from the overall abundance of airborne microbial communities and the aerosol sampler. Airborne bacteria are not as prevalent as their terrestrial or aquatic counterparts. Thus, it was no surprise that the counts were low in our flow cytometry data and taxonomic abundance. However, we collected 35 samples; only 22 samples had enough bacteria to develop libraries we could analyze. Another potential contributor to this discrepancy is the aerosol sampler itself. It is capable of collecting air samples that we can analyze; we have shown this much in our study. However, the device may not be as well suited to the level of detail required for an in-depth dive into microbial ecology. Without a liquid medium within the sampling cone, there is no way to be sure that all bacteria remain in the cone at collection.

Low abundance samples invite the jackpot effect (Lou et al., 2013) to the sequencing procedure. The primers we selected are well suited for a wide range of bacteria; however, if the bacterial DNA is so sparse, it may not bind to any primers during the PCR process. We potentially miss out on characterizing interesting taxa that may have been present in the air. Additionally, the DNA that does polymerize may overestimate the relative abundance purely by chance of arriving at a primer during one of the early reactions.

To mitigate these effects, one would need to investigate the validity of using the CC from a complex ecological experiment by comparing it to a more traditional aerosol sampler. Additionally, running the sampler for perhaps 8 hours instead of 6 could increase the biomass yield during sampling.

Conclusion

Microbial community of the Sea Surface-Atmospheric Boundary Layer (SSABL) is part of an interface where microorganisms from the ocean and the atmosphere interact. The SSABL is a small window between two ecosystems where the microbial community lives in a highly dynamic environment.

We sought to characterize the microbial community of the SSABL, by identifying significant taxa present within the ecosystem, and describing the physical properties that can affect the community structure. The analysis of communities' diversity, absolute abundance, and community make-up suggests that bacterial populations in each environmental type are distinct. Our ordination of these communities highlights two distinct ecological groupings. However, we observed some overlap of airborne samples into the marine group. Suggesting that there are contributing taxa that make atmospheric samples more marine.

Our Study faced challenges primarily due to low abundance of bacterial communities within the atmosphere. Additionally, sequences generated from these samples may not have amplified evenly due to the lack of biomass in our collection process. We suspect that our instrument, the Bertin Coriolis may not be well suited for this level of analysis and can limit the number of microbes collected during the sampling procedure.

Though we faced obstacles in our procedure, we characterized the community by defining taxa present in airborne samples that correlate to a marine-like community. Additionally, we identified several parameters that influence the community structure, shifting the structure to a more marine like ecosystem.

References

Aller, J. Y., Kuznetsova, M. R., Jahns, C. J., & Kemp, P. F. (2005). The sea surface microlayer as a source of viral and bacterial enrichment in marine aerosols. *Journal of Aerosol Science*, *36*(5), 801–812. https://doi.org/10.1016/j.jaerosci.2004.10.012

Amann, R. I., Ludwig, W., & Schleifer, K. H. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological Reviews*, *59*(1), 143–169.

Archer, S. D. J., Lee, K. C., Caruso, T., King-Miaow, K., Harvey, M., Huang, D., Wainwright, B. J., & Pointing, S. B. (2020). Air mass source determines airborne microbial diversity at the ocean–atmosphere interface of the Great Barrier Reef marine ecosystem. *The ISME Journal*, *14*(3), 871–876. https://doi.org/10.1038/s41396-019-0555-0

Bauer, H., Giebl, H., Hitzenberger, R., Kasper-Giebl, A., Reischl, G., Zibuschka, F., & Puxbaum, H. (2003). Airborne bacteria as cloud condensation nuclei. *Journal of Geophysical Research: Atmospheres*, *108*(D21). https://doi.org/10.1029/2003JD003545

Bowman, J. S. (2021). *Scripps Ecological Observatory*. Scripps Ecological Observatory. https://library.ucsd.edu/dc/collection/bb8165246w

Bowman, J. S., & Ducklow, H. W. (2015). Microbial Communities Can Be Described by Metabolic Structure: A General Framework and Application to a Seasonally Variable, Depth-Stratified Microbial Community from the Coastal West Antarctic Peninsula. *PLOS ONE*, *10*(8), e0135868. https://doi.org/10.1371/journal.pone.0135868

Burrows, S. M., Butler, T., Jöckel, P., Tost, H., Kerkweg, A., Pöschl, U., & Lawrence, M. G. (2009). Bacteria in the global atmosphere – Part 2: Modeling of emissions and transport between different ecosystems. *Atmospheric Chemistry and Physics*, *9*(23), 9281–9297. https://doi.org/10.5194/acp-9-9281-2009

Burrows, S. M., Elbert, W., Lawrence, M. G., & Poschl, U. (2009). Bacteria in the global atmosphere – Part 1: Review and synthesis of literature data for different ecosystems. *Atmos. Chem. Phys.*, 18.

Carvalho, E., Sindt, C., Verdier, A., Galan, C., O'Donoghue, L., Parks, S., & Thibaudon, M. (2008). Performance of the Coriolis air sampler, a high-volume aerosol-collection system for quantification of airborne spores and pollen grains. *Aerobiologia*, *24*(4), 191–201. https://doi.org/10.1007/s10453-008-9098-y

Cha, R. S., & Thilly, W. G. (1993). Specificity, efficiency, and fidelity of PCR. *Genome Research*, *3*(3), S18–S29. https://doi.org/10.1101/gr.3.3.S18

Curry, J. A., & Webster, P. J. (Eds.). (1999). Chapter 5—Nucleation and Diffusional Growth. In *International Geophysics* (Vol. 65, pp. 129–158). Academic Press. https://doi.org/10.1016/S0074-6142(99)80027-8

Engel, A., Bange, H. W., Cunliffe, M., Burrows, S. M., Friedrichs, G., Galgani, L., Herrmann, H., Hertkorn, N., Johnson, M., Liss, P. S., Quinn, P. K., Schartau, M., Soloviev, A., Stolle, C., Upstill-Goddard, R. C., van Pinxteren, M., & Zäncker, B. (2017). The Ocean's Vital Skin: Toward an Integrated Understanding of the Sea Surface Microlayer. *Frontiers in Marine Science*, *4*. https://www.frontiersin.org/articles/10.3389/fmars.2017.00165

Fröhlich-Nowoisky, J., Kampf, C. J., Weber, B., Huffman, J. A., Pöhlker, C., Andreae, M. O., Lang-Yona, N., Burrows, S. M., Gunthe, S. S., Elbert, W., Su, H., Hoor, P., Thines, E., Hoffmann, T., Després, V. R., & Pöschl, U. (2016). Bioaerosols in the Earth system: Climate, health, and ecosystem interactions. *Atmospheric Research*, *182*, 346–376. https://doi.org/10.1016/j.atmosres.2016.07.018

Harrison, R. M., Jones, A. M., Biggins, P. D. E., Pomeroy, N., Cox, C. S., Kidd, S. P., Hobman, J. L., Brown, N. L., & Beswick, A. (2005). Climate factors influencing bacterial count in background air samples. *International Journal of Biometeorology*, *49*(3), 167–178. https://doi.org/10.1007/s00484-004-0225-3

Hu, W., Murata, K., Fukuyama, S., Kawai, Y., Oka, E., Uematsu, M., & Zhang, D. (2017). Concentration and Viability of Airborne Bacteria Over the Kuroshio Extension Region in the Northwestern Pacific Ocean: Data From Three Cruises. *Journal of Geophysical Research: Atmospheres*, *122*(23), 12,892-12,905. https://doi.org/10.1002/2017JD027287

Kellogg, C. A., & Griffin, D. W. (2006). Aerobiology and the global transport of desert dust. *Trends in Ecology & Evolution*, *21*(11), 638–644. https://doi.org/10.1016/j.tree.2006.07.004

Legendre, P., Borcard, D., & Peres-Neto, P. R. (2005). Analyzing Beta Diversity: Partitioning the Spatial Variation of Community Composition Data. *Ecological Monographs*, *75*(4), 435–450. https://doi.org/10.1890/05-0549

Lou, D. I., Hussmann, J. A., McBee, R. M., Acevedo, A., Andino, R., Press, W. H., & Sawyer, S. L. (2013). High-throughput DNA sequencing errors are reduced by orders of magnitude using circle sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(49), 19872–19877. https://doi.org/10.1073/pnas.1319590110

Mayol, E., Arrieta, J. M., Jiménez, M. A., Martínez-Asensio, A., Garcias-Bonet, N., Dachs, J., González-Gaya, B., Royer, S.-J., Benítez-Barrios, V. M., Fraile-Nuez, E., & Duarte, C. M. (2017). Long-range transport of airborne microbes over the global tropical and subtropical ocean. *Nature Communications*, *8*(1), 201. https://doi.org/10.1038/s41467-017-00110-9

Non-metric multidimensional scaling—GUSTA ME. (n.d.). Retrieved August 27, 2022, from https://sites.google.com/site/mb3gustame/dissimilarity-based-methods/nmds

Oksanen, J., Blanchet, G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, Peter. R., O'Hara, R. B., Simpson, Gavin. L., Solymos, P., H. Stevens, M. H., Szoecs, E., & Wagner, H. (2020). *vegan: Community Ecology Package. R package version 2.5-7*. McGlinn Lab. https://CRAN.R-project.org/package=vegan

Overview | Southern California Coastal Ocean Observing System. (n.d.). Retrieved August 22, 2022, from https://sccoos.org/overview/

POMEROY, L. R., leB. WILLIAMS, P. J., AZAM, F., & HOBBIE, J. E. (2007). The Microbial Loop. *Oceanography*, *20*(2), 28–33.

Pöschl, U. (2005). Atmospheric Aerosols: Composition, Transformation, Climate and Health Effects. *Angewandte Chemie International Edition*, *44*(46), 7520–7540. https://doi.org/10.1002/anie.200501122

Renaud, F., & Perez, R. (2010). Climate change vulnerability and adaptation assessments. *Sustainability Science*, *5*(2), 155–157. https://doi.org/10.1007/s11625-010-0114-0

Roemmich, D. (1989). Mean transport of mass, heat, salt and nutrients in southern California coastal waters: Implications for primary production and nutrient cycling. *Deep Sea Research Part A. Oceanographic Research Papers*, *36*(9), 1359–1378. https://doi.org/10.1016/0198-0149(89)90088-5

Southern California Coastal Ocean Observing System. (n.d.). Retrieved August 22, 2022, from https://sccoos.org/

Stein, A. F., Draxler, R. R., Rolph, G. D., Stunder, B. J. B., Cohen, M. D., & Ngan, F. (2015). NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System. *Bulletin of the American Meteorological Society*, *96*(12), 2059–2077. https://doi.org/10.1175/BAMS-D-14-00110.1

Terrill, E., Peck, S., Hazard, L., Davis, R. E., DiGiacomo, P. M., Jones, B. H., Keen, C., Moline, M., Orcutt, J., Stolzenbach, K., Washburn, L., Helling, H., Long, J., Magdziarz, S., Laughlin, M., & Kasschau, J. (2006). The Southern California Coastal Ocean Observing System. *OCEANS 2006*, 1–8. https://doi.org/10.1109/OCEANS.2006.306877

Tignat-Perrier, R., Dommergue, A., Thollot, A., Keuschnig, C., Magand, O., Vogel, T. M., & Larose, C. (2019). Global airborne microbial communities controlled by surrounding landscapes and wind conditions. *Scientific Reports*, *9*(1), 14441. https://doi.org/10.1038/s41598-019-51073-4

Walters, W. (2016). *Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys*. https://doi.org/10.1128/mSystems.00009-15

Welsh, D. T. (2000). Ecological significance of compatible solute accumulation by micro-organisms: From single cells to global climate. *FEMS Microbiology Reviews*, *24*(3), 263–290. https://doi.org/10.1111/j.1574-6976.2000.tb00542.x

Wilson, T. W., Ladino, L. A., Alpert, P. A., Breckels, M. N., Brooks, I. M., Browse, J., Burrows, S. M., Carslaw, K. S., Huffman, J. A., Judd, C., Kilthau, W. P., Mason, R. H., McFiggans, G., Miller, L. A., Nájera, J. J., Polishchuk, E., Rae, S., Schiller, C. L., Si, M., … Murray, B. J. (2015). A marine biogenic source of atmospheric ice-nucleating particles. *Nature*, *525*(7568), 234–238. https://doi.org/10.1038/nature14986

Wiśniewska, K., Lewandowska, A. U., & Śliwińska-Wilczewska, S. (2019). The importance of cyanobacteria and microalgae present in aerosols to human health and the environment – Review study. *Environment International*, *131*, 104964. https://doi.org/10.1016/j.envint.2019.104964

Xia, X., Wang, J., Ji, J., Zhang, J., Chen, L., & Zhang, R. (2015). Bacterial Communities in Marine Aerosols Revealed by 454 Pyrosequencing of the 16S rRNA Gene. *Journal of the Atmospheric Sciences*, *72*(8), 2997–3008. https://doi.org/10.1175/JAS-D-15-0008.1