Exploring eDNA methodologies as a way to detect relationship between Anchovies and Krill in Monterey Bay Canyon

Bryce Corbett, Southern Illinois University-Carbondale

Mentor: Francisco Chavez
Summer 2016

Keywords: eDNA, krill, Monterey Bay, mtCO1

ABSTRACT

My research used environmental DNA (eDNA) to look at the distribution of zooplankton in Monterey Bay, California. eDNA is material that is sloughed by animals and does not require sampling live or whole organisms. In the fall months of 2013 through 2015 there was an increase in whale sightings and their prey, anchovies in Monterey Bay Canyon relative to the previous years (2010-2012). Was this increase in anchovies and subsequent whale sightings related to an increase in zooplankton? We hypothesized that zooplankton DNA would be present in large amounts in the environmental samples collected during 2013-2015. The second hypothesis is that zooplankton are less abundant over the shallow shelf than over deeper the slope. Zooplankton vertically migrate to escape predation and are not able to migrate deeply over the shelf. DNA was extracted from water samples procured at C1 Coastal Station in Monterey Bay during the fall months of 2010 through 2015 for the anchovy and whale study. Similar extractions were performed from a cruise in Monterey Bay in October 2015 when samples were collected over the shelf and slope. Polymerase Chain Reaction (PCR) was used to amplify the extracted DNA for the mitochondrial cytochrome oxidase gene, targeting invertebrates. After receiving the correct amplicons a PCR cleanup was done and the final target DNA was sent to Stanford University to be sequenced by MiSeq.

INTRODUCTION

Climate change is the primary cause of habitat disruption in the ocean. Since the industrial revolution CO_2 concentrations have increased by 40% (Shuang-Jing et.al 2014). Half of that amount is absorbed by the world's oceans resulting in acidification or the lowering of the ocean's pH. This causes problem for marine organisms by disrupting the ocean carbon cycle. As atmospheric CO_2 levels rise this increases the effect of greenhouses gases that trap heat. An increase in surface temperature can cause a negative reaction in lower trophic level organisms such as zooplankton. When lower level organisms and their ability to acquire energy is effected it has a domino effect, harming the largest of animals such as humpback whales. Climate change is harming the species diversity in the world's oceans.

Knowledge of life in the open ocean is limited by a variety of factors. Both spatially and temporally it is difficult to keep a consistent surveying presence of different communities, while poor visibility and rough conditions only add to the difficulties. Further, it is difficult to quantify organisms that are rarely seen and to observe their feeding or reproductive behaviors. In addition, the equipment needed to tag or sample these populations can be expensive and disruptive to their behaviors. However; monitoring biodiversity in the ocean is essential. Climate change and human activities are thought to have largely negative effects on the ocean and the organisms living in this habitat. To conserve aquatic species, it is necessary to know how species react to this change and how they can resist negative responses. Identifying where specific species are concentrated for sampling and observation is important for conservation. Analyzing (eDNA) is an emerging method that is relatively inexpensive and has low impact on target populations. These factors give eDNA the potential to be used as a powerful ecological survey tool.

eDNA is the DNA left behind by multicellular organisms in the form of waste, saliva, skin cells, etc. (Davy, Kidd & Wilson 2015). In this project we will use eDNA to test the relative presence or absence of krill, a type of zooplankton, in Monterey Bay from 2010 to 2015. Humpback whale sightings were more frequent after 2012. This increase in sightings is most likely due to increased concentration of anchovies, the favorite prey for humpbacks. Anchovies feed on krill and increased concentrations of this organism may explain the high abundance of anchovies during 2013, 2014 and 2015. Habitat compression, a phenomenon in which an organism is displaced from its original habitat, may be another reason why anchovies were

concentrated near shore close to Moss Landing. With climate change and the subsequent increase of sea surface temperature, dissolved oxygen (DO) decreases. Low DO causes metabolic stress on marine organisms, forcing them to spend more energy on anaerobic processes (Prince & Goodyear 2006). To escape this stress, organisms will migrate to more habitable areas.

The relationship between anchovies and krill can be reduced to a simple predator-prey interaction; the anchovies were present to feed on the krill. Using eDNA methods we can confirm this conclusion. eDNA offers the ability to monitor biodiversity without affecting the target population is a powerful tool.

In order to use the eDNA gathered from the water samples, genetic barcodes are necessary. DNA barcoding uses short sequences to identify species in a particular habitat. It allows the identification of known species that may be difficult to observe or very fragile (Bucklin, Steinke, & Blanco-Bercial 2011). Mitochondrial cytochrome c oxidase subunit 1 (mtCO1) is a maternally coded gene used for species identification and has various advantages over the nuclear genome. Those advantages include, mtCO1 lacking introns, its limited exposure to recombination and the high copy numbers in every cell. In addition mtCO1 evolution is rapid enough that it allows differentiation between closely related species.

MATERIALS AND METHODS

Samples were taken from two different pools of data: the Canon15 (controlled agile & novel observing network) cruise in October 2015, and historic samples acquired during the fall months of 2010 through 2015. The Canon15 was chosen for sampling based on the data gathered from the EK500 echosounder (fig. 2).

During October 2015 there were large amounts of krill aggregates in the Monterey Bay Canyon area. The echosounder output two different graphs with different frequencies: one at 38 kHz the other at 200 kHz. Water samples to be used for eDNA analysis were collected at 9 different sites. 3 samples were taken from the shelf of the canyon, another 3 from sites on the slope of the canyon and the final 3 from offshore sites (fig. 3). Bottles 12 and 7 from the CTD were used; 12 collects water from the surface and 7 collects water from a depth of approximately 40 meters. The goal of collecting samples from various sites in the canyon was to compare zooplankton composition at those stations.

WATER SAMPLE COLLECTION

Dense collections of zooplankton were found using the EK500 echosounder data in October 2015. eDNA samples were taken at the site of these aggregations at surface depth >5m. Conductivity, temperature and depth devices (CTD) were used to collect at various depths (see table 1). CTD rosette samples were collected from 11 depths at 3 locations in Monterey Bay: C1 coastal station, Mooring1 mid-bay station, and Mooring2 outer bay station. Samples were collected each year, once a month in August, September and October. They were filtered using a .22um filter in labeled test tubes and stored in liquid nitrogen to preserve target samples.

DNA EXTRACTION AND AMPLIFICATION

Our methods of amplification and extraction were adapted from methods by Kelly et.a al. DNA was extracted using Buffers ATL, AW1, AW2 & proteinase K. This step involved heavy use of 10-100ul pipettes and various specialized test tubes to specifically aid in DNA extraction. We ran samples through a vacuum manifold and washed each twice with ethanol and once with AW1 buffer. After vacuuming, the samples were spun for three minutes at 14,000 rotations per minute in a centrifuge to dry them. We eluted DNA using the centrifuge and an AE buffer, stabilizing the DNA within the sample for long-term storage.

Following extraction, we used NanoDrop, a type of UV-Vis spectrophotometer, to assess the quality and concentration of DNA. We used PCR as a tool to amplify mtCO1 in triplicate reactions. We mixed samples with a forward and reverse primer made for the amplification of mtCO1. In addition to these primers, bovine serum albumin (BSA) was added as an inhibitor, preventing non-target DNA from being amplified. To verify amplification of the correct genes we placed samples in a gel electrophoresis plate. Gel electrophoresis uses agar gel and electric currents to separate mixtures of DNA dependent on base pair length.

AMPure beads function as adhesive surfaces for DNA to stick to. These beads are added to the PCR product plate and placed on a super magnet. The beads are washed twice with ethanol and left to dry. The washing should rid the DNA of the BSA, primers and Mastermix added prior. After beads have dried TE buffer is added to elute the DNA off the beads. We will then

pipette DNA into another gel electrophoresis plate. The results should be pure DNA without primers.

LIBRARY POOLING

240 ng of DNA from each sample was transferred into a sterile tube.The samples will be combined in a single gene library pool and sent to Stanford with sequencing primer. This DNA will be put through a fluorometer that will measure the concentration of DNA in each sample in nanograms per liter. Stanford will sequence the DNA using an Illumina Miseq instrument that will output all the DNA sequences extracted from our samples.

DATA PIPELINE

This information will then be entered into the data pipeline on an MBARI server. The data pipeline consists of several program all performing different tasks. These programs include: Paired end sequence reads, the Paired-End reAd mergeR or PEAR , quality filters. swarm, seqtk, python, blast+, MEGAN, R, awk, and sed. Pair end sequence reads take the 5' and 3' end of DNA and generates high quality alignable sequence data (citation here). PEAR generates reads from both ends of target DNA and works when fragment lengths are varied. PEAR is also advantageous because it does not require previous information on read length nor target fragment size (Zhang et. al. 2014). Quality filters make the target fragments easier to analyze and can increase accuracy of the final output. These filters include cutadapt and vsearch. Cutadapt finds and removes parts of the DNA sequences d such as poly-A tails, primers and adapter sequences (Martin 2011). When Miseq finishes its run, the output is dozens of zipped .fastq/ .fasta files per each sample. Seqtk can read and process these files. Python is a simple programming language that is just used to combine duplicate sequences if found in the fastq/a files. The Basic Local Alignment Search Tool or BLAST is an algorithm that identifies similarity between sequences. BLAST assembles segments of DNA into discrete comparable parts that are used by MEGAN (MetaGenome Analyzer). MEGAN is another taxonomic assignment tool that uses the BLAST output and compares resulting sequences to gene sequences in GENBANK. R is an ecological analysis tool that produced a variety of graphs indicating species diversity and abundance of DNA reads. Awk is a programming language that processes patterns within text. Each program is connected to the next and requires parameters from previous programs to run. The end result will

be a taxonomic table produced from target DNA, with identification either partially sequenced to family name or fully sequenced to species name.

RESULTS

Samples from CANON15 and the time series from fall months of 2010 to 2015 yielded little to no krill DNA reads. Using BLAST and operational taxonomic unit tables we were able to identify the species that were present in the sample. The only Euphausia species identified was E. *pacifica.* The species that were in highest abundance were marine copepods such as Cyclopoida and *Thermocyclops inversus* . There was a small amount of *Euphausia pacifica* DNA represented, which was expected because it is one of the most common krill species in Monterey Bay. While we did not find large amounts of krill DNA we did find a multitude of organisms and patterns within our data. For example, genetic evidence was found of the pelagic red crab at stations from the historical samples taken during the years 2011 and 2013. It is expected that these crabs are present during El Nino years which occurred in 2015, however they were not expected during non El Nino years of 2011 and 2013. Additionally, we discovered that communities were similar across depths regardless of distance. In the CANON15 samples surface samples that were taken 40 km away from each other exhibited the same makeup of organisms. This pattern was also reflected in samples taken 40m deep (fig. 4).

DISCUSSION

My results have a variety of implications; one is that krill were not in high amounts during the years of high anchovy concentration or two, that eDNA methods are not currently successful means of identifying krill DNA in the Monterey Bay Canyon. However there is a third option that came up after surveying the taxonomic we did receive. We found high amounts of Scatella and Sciaridae which are two genuses of flies. Scatella are flies that are often found on the greenhouses, however we found DNA in offshore stations. Sciaridae is a genus commonly referred to as dark winged fungus gnats that occupy houseplants or more inland areas. The fact these flies were found in such high abundance indicates that there could have been misidentification of gene reads. Krill DNA could have been more abundant than our results

show, but the fragments were too small for the library repositories to assign them as krill, instead giving them the taxonomic name of the next closest thing, i.e. flies.

eDNA methods are very new and include many detailed steps, because of this there are many opportunities for errors. In addition, DNA is a highly sensitive substance and can degrade quickly if not stored properly.

This research will demonstrate how gene sequencing can be used to identify presence of marine organisms without visual confirmation. As the effects of anthropogenic climate change become more apparent, the need for greater understanding of the ocean's populations is increasing. eDNA has the potential to greatly increase ecological monitoring of the oceans. By using eDNA researchers will not have to remove an organism from its habitat to confirm its presence. This will reduce the disturbance biodiversity surveys have on their target populations. eDNA's effectiveness at determining which organisms were present in a specific area are being tested. While eDNA methods have potential there are many drawbacks. These include the fact that these methods are very new and include many detailed steps. Due to this, there are many opportunities for errors. For example, during DNA extraction many things could go wrong. DNA is a highly sensitive substance that can degrade quickly if not stored properly. In addition, in the water column DNA lasts for about 3 days after a marine organism has left it behind. This means that collection eDNA must be quick and done with purpose. eDNA is not abundance based, the number of organisms cannot be determined by eDNA methods alone, only their presence.

MtCO1 as a barcoding gene was successfully used to identify various taxonomic groups. However; sponges and anthozoans are more difficult to identify using mtCO1. The anthozoans and poriferans have mtCO1 sequences that have not diverged enough to be differentiated (Bucklin 2011). The anthozoa class contains coral species which are some of the most ecologically important groups in the marine ecosystem. As we continue our use of genetic barcodes such as mtCO1 to sequence eDNA, the possibility of finding barcodes that can identify these species increases.

CONCLUSIONS

Organizations like the Marine Biodiversity Observation Network (MBON), a coalition of scientists whose purpose is to study the ocean and construct management strategies informed by

the data they gather, are striving to improve their sampling methods continuously. To evaluate the open sea habitat, MBON must be able to identify organisms that live there and the size of their populations and eDNA could do just that. It is important for the scientific community to consistently come up with cost effective ways to observe the world around us.

The changes seen in the ocean as a result of anthropogenic activity will have some of the biggest impacts worldwide. The global population currently consumes 110-130 million tonnes of seafood annually (Steiner 2009). The centrality of ocean resources in the global diet and economy make the need for large scale conservation efforts quite clear. Sea surface temperature and biodiversity have an inverse relationship, as SST increases, species diversity decreases. While using the time series of fall 2010 to fall 2015 we reported this inverse relationship on a graph (fig.5). While there was a decrease in species diversity in 2013, 2014 and 2015 there was still a large amount of anchovies and whale sightings. This could be the result of the phenomenon mentioned earlier, habitat compression. The anchovies were displaced from their natural habitat from increased SST and the humpback whales followed their movement towards the mouth of the Monterey Bay Canyon.

eDNA is one way to monitor biodiversity and give quantitative evidence of the effect climate change has on the ocean environment. Finding the link between species diversity and the effect climate change has on that diversity is integral to finding a solution.

ACKNOWLEDGEMENTS

I would like to extend a thank you to Francisco Chavez, my mentor, for introducing me to eDNA methods; Kris Walz for walking me through DNA extraction and being extremely patient. Reiko Michisaki for producing many species diversity graphs and making sure I got my end product full of taxonomic names. MBARI for hosting me as an intern, specifically George Matsumoto and Linda Kuhnz for being helpful sources of support. The CSUMB REU program, without which i would not be able to do my work at MBARI. Finally, the NSF for funding this project.

References

Bucklin, A., Steinke, D., & Blanco-Bercial, L. (2011). DNA Barcoding of marine Metazoa. Annual Review of Marine Science Annu. Rev. Marine. Sci., 3(1), 471-508.
doi: 10.1146/annurev-marine-120308-080950

Bucklin, A., Lindeque, P.K., Rodriguez-Ezpeleta, N., Albaina, A. & Lehtiniemi, M. (2016). Metabarcoding of marine zooplankton: prospects, progress and pitfalls. Journal of Plankton Research., 38(3), 393-400.
doi: 10.1093/plankt/fbw023

Davy, C.M, Kidd, A., Wilson, C.C (2015). Development and Validation of Environmental DNA (eDNA) Markers for Detection of Freshwater Turtles. PLoS ONE., 10(7): e0130965.
doi:10.1371/journal.pone.0130965

Kelly, R.P., Port, J.A., Yamahara K., & Crowder, L. (2014). Using Environmental DNA to Census Marine Fishes in a Large Mesocosm. PLoS ONE., 9(1): e86175.
doi:10.137/journal.pone. 0086175

Martin, M. (2011). Cutadapt removes Adapter Sequences from High-Throuput Sequencing Reads. EMBnet. 17(1), 10-12.
Doi: 10.14806/ej.17.1.200.

Prince, E.D., Goodyear, C.P, (2006). Hypoxia-based habitat compression of tropical pelagic fishes. Fisheries Oceanography.,15(6), 451-464

Shuang-Jing, W., Long, C., Na, L., (2014). Responses of the ocean carbon cycle to climate change: Results from an earth system climate model simulation. Advances in Climate Change Research. ScienceDirect., 5(3), 123-130.
doi: 10.1016/j.accre.2014.11.004.

Steiner, A.(2009). WORLD FOOD SUPPLY. GRID-Arendal. UNEP.

Zhang, J., Kober, K., Flouri, T., Stamatakis, A. (2014). PEARL a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics. 30(5). 614-620.
doi: 10.1093/bioinformatics/btt593

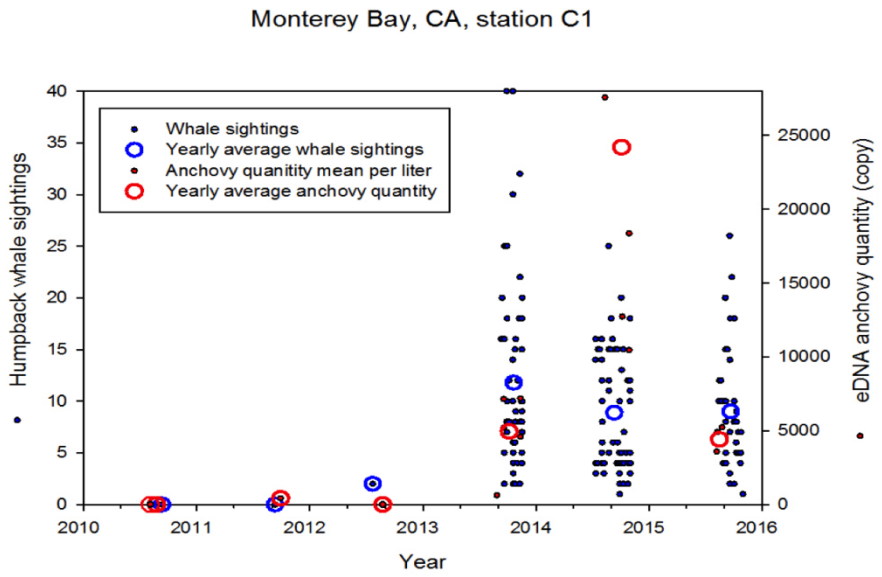
FIGURES & TABLES

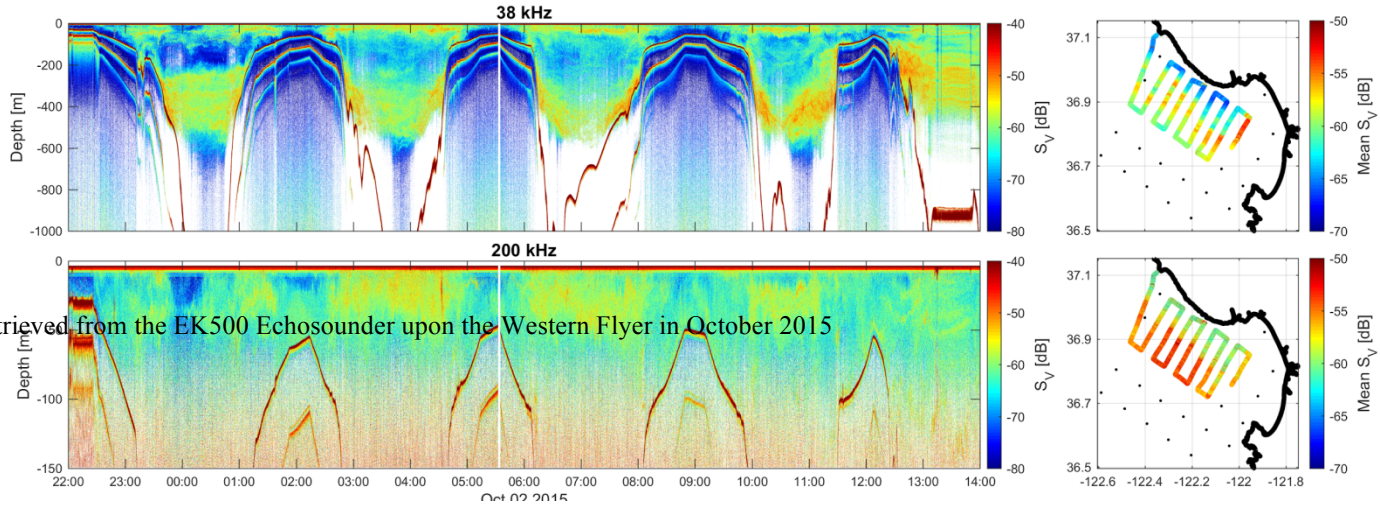
Table 1: Depicts the sample ID and depth at which each sample was taken. All samples were collected from Coastal station 1 located on the edge of the Monterey Bay Canyon shelf

Samples	Location	Depth(m)
28512c01_12_hplc	C1 Coastal Station (C1)	1.885
23208c01_11_hplc	C1	1.319
26208c01_11_hplc	C1	1.241
23009c01_11_hplc	C1	1.078
25209c01_11_hplc	C1	1.775
21410c01_11_hplc	C1	1.34
23810c01_11_hplc	C1	1.307
s410c01_12_hplc	C1	0.969
21511c01_12_hplc	C1	2.062
23611c01_12_hplc	C1	2.032
25511c01_12_hplc	C1	2.012
canon11c01_12_hplc	C1	5.387
C0912c01_12_hplc	C1	2.017
22112c01_12_hplc	C1	1.889
23512c01_12_hplc	C1	2.354
CN13IDc01_12_hplc	C1	0.677

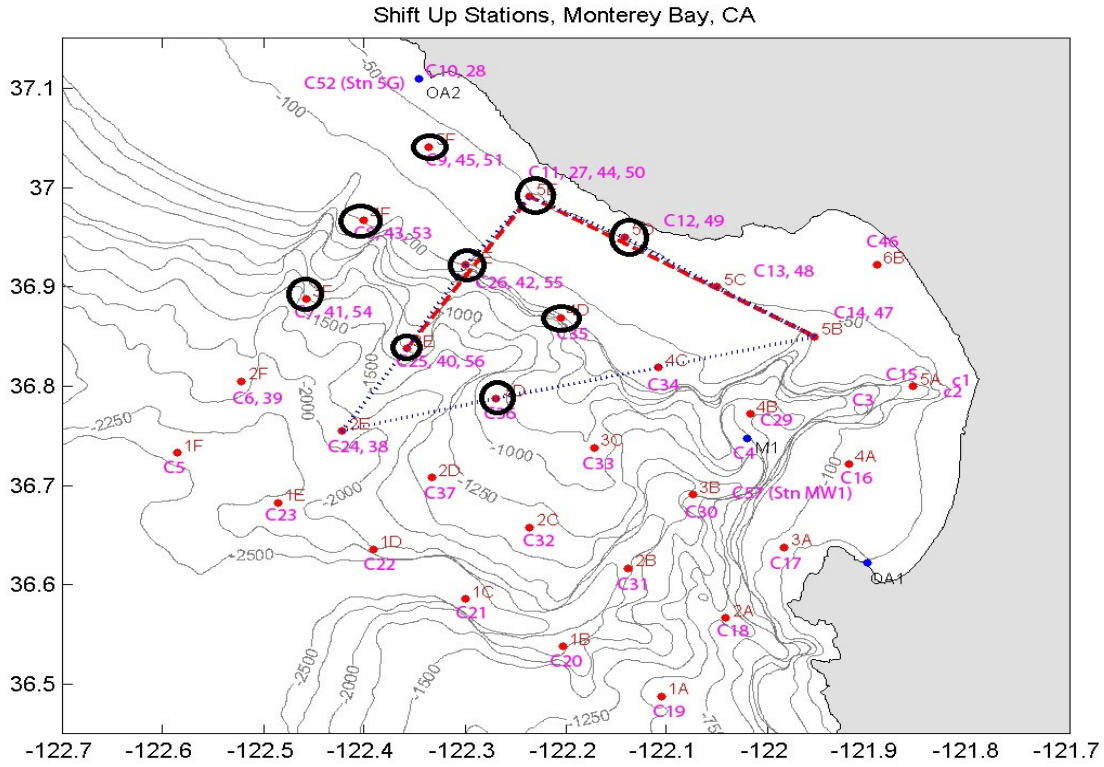
ıths of 2010 through 2015

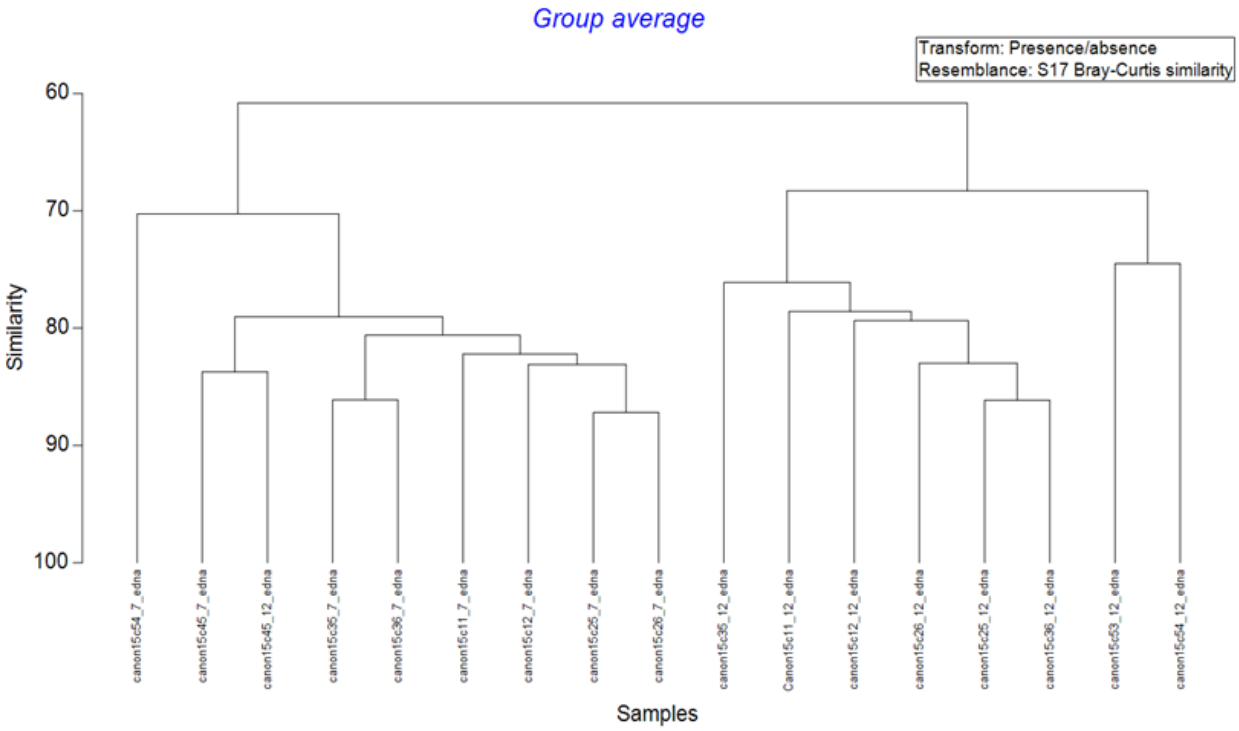
22013c01_12_hplc	C1	0.947
24013c01_12_hplc	C1	1.963
21515c01_12_hplc	C1	1.805
22414c01_12_hplc	C1	1.818
23715c01_12_hplc	C1	1.652
canon13c01_12_hplc	C1	0.666
canon13c01_12_Astar	C1	0 .666
31613c01_12_hplc	C1	1.428
28014c01_12_hplc	C1	1.047
30214c01_12_hplc	C1	1.187

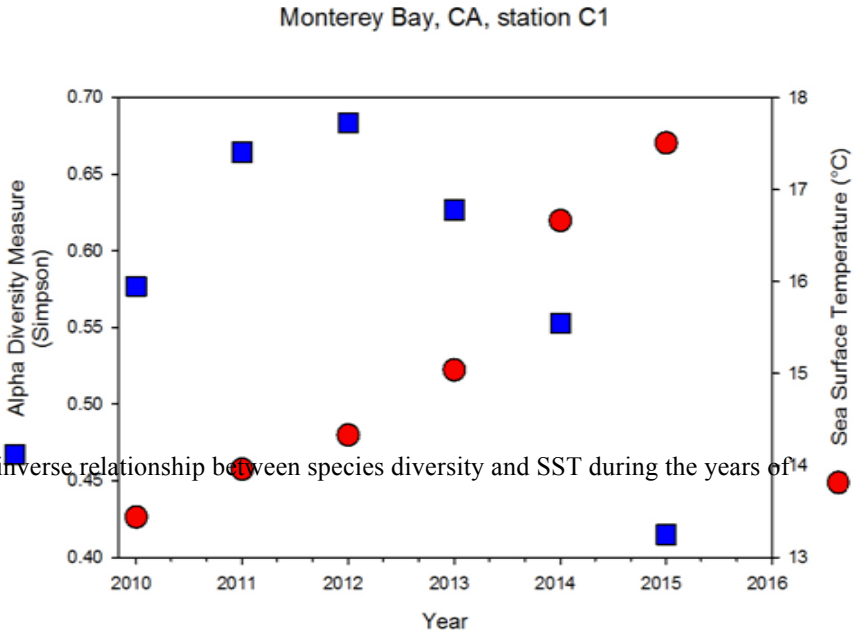




2 Data retrieved from the EK500 Echosounder upon the Western Flyer in October 2015







re 5 Graph indicates the inverse relationship between species diversity and SST during the years of to 2015