EcologicalGenomics\_Homework1

Bailey Kretzler

9/29/2021

## Background

Sunflower sea stars (*Pycnopodia helianthoides*) are important predators in coastal ecosystems of western North America. In recent years, sunflower sea star populations have been significantly impacted by sea star wasting diseases (SSWD). Up until 2016, SSWD had not been observed in the species most northern range. However an increase in SSWD events in Alaska was noticed by the Alaska department of fish and game. As such, a sampling effort was undertaken to assess the microbiomes of sea stars at recently impacted and naive sites. Epidermal samples were collected from seven sites across south eastern Alaska, from which RNA was extracted and the V3/V4 16s region was sequences. Previous studies have found that the microbiome of sick and health sea stars differ (melissa’s paper). Still is unclear if differences in the microbiome of individuals at naive and impacted sites differ. As a result, I hope to examine the following questions: **1) Are the microbiomes of healthy individuals at sick sites more similar to that of healthy individuals at healthy sites or sick individuals at sick sites? What specific taxa are in greater abundance at each site and what are their potential functions in seastar microbiomes?**  To test this we can compare the abundances of key taxa from the individuals that are healthy at naive sites (HH), healthy at impacted sites (HS) and sick at impactes sites (SS). **I hypothesize that the HS indviduals will resemble an intermediate between HH and SS individuals, where they have generally healthy and beneficial microbiomes with a rise in commensalist and opporunists. It is also possible that pathogenic microbes will also be present that are outcompete by the beneficials and commensalists.**

## Bioinformatics Pipeline

Once sequence reads were obtained, I used QIIME2 (Boylen et.al. 2019) to denoise and analyze the data. The first step was to denoise the paired end reads by trimming and truncating both the forward and reverse reads(supplemental code 1). Denoising the reads allows us to work with a cleaner set of data, which is prone to fewer misread basepairs (bp) per the total. After viewing a summary plot for read quality cut offs were set at 16bp forward trim, 0bp reverse trim, 278bp forward truncate, and 220bp for reverse truncate. These cut off were determined using a phred score cut off of 20 representing 99% basepair accuracy. Once data were denoised, the resulting sequences, metadata and denoising statisitcs were extracted for viewing(supplemental code 2). The sequences and metadata were used for further data analysis.

Once denoised sequences were obtained, a phylogenetic tree was constructed to determine the identities and relationships among taxa (supplemental code 3). The phylogenetic tree was then used to determine core diversity metrics such as the shannon diversity index and the unweighted and weighted unifrac distance (supplemental code 4). To calculate core diversity metrics a sampling depth of 9000 reads was used, based on the number of sequence features retained vs lost at this depth. Core diversity metrics were then used make visualizations for weighted unifrac distance between site-animal health pairings. Taxa were also tested for differences in abundance by comparing the sequence reads from the sea star microbiome to sequences of known classification via the greengenes database (supplemental code 5). The taxa classification was then used to develop a bar plot comparing the abundance of each taxa across the different site-animal conditions (HH,HS,SS).

## Results

Across the 85 samples retained in the analysis, a total of 10,796 features were identified. The frequency per sample of these features ranged from 18,751 to 154,925. The total number of forward and reverse read for the samples was 24,256,727 with a minimum and maximum reads per sample of 94,407 and 438,074 respectively. From the weighted unifrac beta diversity estimate, it is clear that SH is more similar to HH that to SS (figure 1). Further, site-animal health pairings differ in abundance for 119 taxa (supplemental code 6). As suspected SH sea stars do not identically resemble the HH or SS individuals but rather some type of intermediate. This is evident in the relative frequency of an unknown genus in the *Spirochaetaceae* and *Flammeovirgaceae* families and the genus *Vibrio* in the *Vibrionaceae* family. The frequency of the *Vibrio* genus significantly increases from HH and SH individuals, with the frequency in SH individuals being similar to that of SS individuals (figure 2, p < 0.001). Alternatively the frequency of the unknown *Spirochaetaceae* and *Flammeovirgaceae* genera decreases from HH to SH individuals, but is still greater than that of SS individuals (figure 2, p < 0.001). Further there is an apparent increase in genus *Psychrilyobacter* in SS individuals which is not seen in HH or SH sea stars (figure 2, p < 0.001). In SS individuals where there is not a large increase in *Psychrilyobacter* there seems to be a comparable increase in the order *Bacteriodales* (figure 2, p < 0.001). Lastly it is worth noting that genus *Psuedoaltermonas* shows the highest frequency in SH sea stars (figure 2, p < 0.001).

## Conclusions

From the data presented it is clear that taxa differ in abundance between HH, SH, and SH sea stars. In some cases, the abundance of specific taxa in SH individuals seems “between” that of HH and SS individuals. SH displaying intermediate levels of abundance is best seen in the unknow genera of families *Spirochaetaceae* and *Flammeovirgaceae* and genus *Vibrionaceae vibrio*. This gradient could suggest the early stages of disease for SH sea stars or a priming effect on sea star immune systems that enable resilience to SSWD. Genera *Psychrilyobacter* and order *Bacteriodales* are suspect as potential pathogens to the sunflower sea star, as either overt or opportunist pathogens. It is important to note that contradictory results have been found in ochre sea starts where *Pseudoaltermonas* was more abundant in healthy seastars, suggesting a divergent role in either the sea star species or environment (LLoyd and Pespeni 2018).

It should be considered that the functions of these species are unknown, making it hard to assume their roles or impact on the sea star microbiome. Further, unequal replication between site-animal health pairings and variation in environmental factors at each site should be considered when examining the data. The qiime2 interface with r made some more robust visualizations difficult but overall the methodology was straight forward and appropriate for the data set. Future work should be conducted to see how environmental conditions, regardless of disease infection, are impacting the presence/absence of certain bacteria and the plasticity of microbiomes across environments or under abiotic stress conditions. Further, the impact of food source should be considered as a factor influencing the sea star microbiome. Lastly, co-occurence analysis of the various taxa would be useful for understanding if the presence of certain microbes is consistently correlated to that of others.

### Citations:

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu YX, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, and Caporaso JG. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology 37: 852–857. <https://doi.org/10.1038/s41587-019-0209-9>

Lloyd, Melanie M, and Melissa H Pespeni. “Microbiome shifts with onset and progression of Sea Star Wasting Disease revealed through time course sampling.” Scientific reports vol. 8,1 16476. 7 Nov. 2018, <doi:10.1038/s41598-018-34697-w>