Introduction

Seagrasses provide several critical marine ecosystem services, including provisioning of food and habitat and acting as an indicator for the overall health of the ecosystems in which they are present (Lamb et al., 2017; Hogarth, 2015). Seagrasses also facilitate ecosystem health by trapping fine sediments and bacteria that are introduced by storm runoff or resuspended from the seafloor. By trapping sediments, eelgrasses can have a direct impact on pollutants, aiding in increased water clarity and purity (Gacia and Duarte, 2000; Lamb et al., 2017). Zostera marina and Zostera pacifica are species of seagrass, commonly referred to as eelgrasses, which naturally occur in the National Marine Fisheries Service (NMFS) West Coast Region. Included in this region are California, Washington, and Oregon (NOAA, 2018). Within California, these eelgrasses may be found along the entire coast from Imperial Beach in Southern California to Humboldt County in the North, but they are concentrated within the San Francisco, Humboldt, and San Diego Bays (Merkel & Associates, Inc., 2014).

Previously, the Z. marina microbiome has been studied by Ettinger et al. (2017) in Bodega Bay, California, where researchers examined the microbial communities found in association with different components of the eelgrass bed, including roots, leaves, and sediments. Using 16S rRNA gene sequencing, they were able to identify significant differences between the microbial communities of each sample type, i.e., between samples taken from the roots vs. leaves vs. sediments. Ettinger et al. (2017) also found that microbial communities differed between samples taken from varying distances to the edge of the seagrass patch, but only among the sediment samples. They did not observe this pattern between samples of the leaf and root microbiomes. The researchers examined how microbial communities differed from sediment found within the eelgrass patches compared to sediment from the edge of the eelgrass patches and outside of the patches; they found significant differences in overall microbial composition, with the most abundant orders being Bacteroidales (highest inside the patch), Flavobacteriales (highest on the patch edge), Desulfovibionales (highest inside the patch), Thioc八erales (highest outside of the patch), Clostridiales (highest inside the patch) and Alteromonadales (highest
outside of the patch). Although the exact mechanism for this phenomenon is uncertain, they observed a correlation between microbial community structure, sampling location, C:N ratio, and eelgrass density.

The abilities of *Z. marina* to reduce pathogen load and alter microbial community structure have also been highlighted in previous work. *Enterococcus* assays showed *Z. marina* had a significant impact on bacteria in eelgrass meadow ecosystems, reducing viable *Enterococcus*, an indicator of fecal contamination (Ortega et al., 2009). Lamb et al. (2017) concluded that *Z. marina* eelgrass meadows found in the intertidal regions of islands in the Spermonde Archipelago, Indonesia, significantly reduced bacterial and pathogenic load in the surrounding waters to the benefit of both humans and marine organisms. The 16S rRNA gene sequencing and *Enterococcus* assays suggested a 50% reduction in the abundance of animal and human pathogens when eelgrass meadows were present, along with a two-fold reduction in coral disease in adjacent reefs (Lamb et al., 2017).

Although the specific mechanisms by which seagrasses influence microbial community structure are not well constrained, some seagrasses have been found to exhibit antimicrobial properties both in the ocean and *in vitro* (Alam et al., 1994; Choi et al., 2009; Mayavu et al., 2009; Kannan et al., 2010). For instance, methanol extracts from the eelgrass *Enhalus acoroides* exhibited antibacterial behavior on Gram-positive bacteria and the model pathogen *Pseudomonas aeruginosa* (Alam et al., 1994). Along with antibacterial potential, extracts from *Z. marina* have also demonstrated antioxidant properties (Choi et al., 2009). These experiments suggest that eelgrasses may directly influence microbial community structure in the coastal marine environment, including reducing the pathogen load.

In San Diego Bay, untreated sewage and storm runoff can influence microbial community structure and introduce pathogenic bacteria. San Diego coastal waters from Imperial Beach to as far north as Coronado are periodically exposed to sewage from the transboundary Tijuana River. During rain events, the Comisión Internacional de Límites y Agua (CILA) in Tijuana shuts down its pump station sewage diverter, resulting in the subsequent discharge of sewage into the Tijuana River estuary (U.S. Department of Justice, 2018). This water contains a variety of sewage and trash, including microbial contaminants, and can result in the closure of public beach access (U.S. Department of Justice, 2018; San Diego County, 2018). San Diego Bay is also periodically exposed to toxic stormwater runoff from Chollas Creek (Schiff et al., 2003; US Department of Justice, 2018). Furthermore, the bay has a history of locally sourced industrial pollution. In the decades preceding 1960, San Diego Bay was exposed to a constant influx of raw sewage and untreated discharge from a variety of industries, including fish canneries, kelp processing facilities, aircraft manufacturing plants, commercial shipyards, and multiple U.S. Naval installations (Fairey et al., 1998). More recently, heavy urban development in San Diego Bay and cross-border sewage outflow have raised concerns about water quality (San Diego County, 2018).

In this study we used 16S rRNA gene sequencing and flow cytometry to test whether local eelgrass meadows influence the structure of the marine microbial community in San Diego Bay, and whether this influence includes a reduced pathogen load. We examined the structure of bacterial and archaeal assemblages from water samples collected from paired sites where eelgrass beds are either present or absent. Sample pairs were collected from two areas, within and just outside of the San Diego Bay.

**Methods**

**Sample collection**

Sites likely to contain eelgrass were identified from the 2014 San Diego Bay Eelgrass Inventory (Merkel & Associates, Inc., 2014), the most recent survey available. The presence of eelgrass was confirmed visually; however, mapping the extent of eelgrass meadows was beyond the scope of this study. Water samples at eelgrass-present and eelgrass-absent sites were collected from a small boat via a peristaltic pump on 3 July 2017 (Figure 1, Table S1). Water was collected from either 10 cm (estimated) off the seabed, or in the mid-water column as indicated in Table S1. In the latter case the point of sample collection was less than 2 m above the seabed. For DNA samples, approximately 1 L of water was filtered through a sterile 0.2 µm Supor filter (Pall), and the filter was immediately placed on dry ice with longer-term storage at –80°C. Flow cytometry samples were collected into sterile 15-ml falcon tubes and immediately placed on ice. Chlorophyll *a* concentration, yield, and phytoplankton photosynthetic efficiency (*F*0/*F*∞) were measured with an AquaFlash handheld active fluorometer (Turner Designs) following the manufacturer’s instructions. Temperature, salinity, and turbidity were measured with a YSI ProDss (Xylem).

**DNA extraction and sequencing**

DNA was extracted with the DNEasy PowerWater DNA extraction kit (Qiagen) within two weeks of sampling. Extracted DNA was quantified using the Qubit HS DNA quantification kit (Invitrogen) and then quality-checked by gel electrophoresis and PCR amplification of the 16S rRNA gene using primers 515F and 806R (Walters et al., 2015). High quality extracted DNA was submitted to the Argonne National Laboratory sequencing center for amplification and library preparation with the same primer set, then paired-end sequenced on the Illumina MiSeq platform. Sequence reads and accompanying metadata were uploaded to NCBI SRA under BioProject PRJNA483963.

**Flow cytometry**

Samples for flow cytometry were stored in the dark at 4°C, and fixed to a final concentration of 4% formaldehyde within 12 hours of sampling. Fixed samples were stained with SybrGreen I (Molecular Probes), spiked with a known concentration of 123-count beads (ThermoFisher), and quantified on a Sysmex CyFlow Space flow cytometer. High nucleic acid (HNA), low nucleic acid (LNA), and overall bacterial cell counts were determined from
populations identified from the forward scatter (FSC) and 488 nm (FL1) parameters using a self-organizing map (SOM) following Bowman et al. (2017).

**Bioinformatics**

Illumina MiSeq reads were demultiplexed using the ‘iudemultiplex’ command in IlluminaUtils (Eren et al., 2013). Demultiplexed reads were quality-controlled and denoised using the ‘FilterandTrim’ and ‘dada’ commands within the R package dada2 (Callahan et al., 2016), and assembled with the ‘mergePairs’ command. The final, merged reads all had mean quality scores ≥ 30. The R package dada2 is designed to denoise 16S rRNA amplicon data such that the final “unique” reads reflect true biological diversity. The non-redundant fasta files of unique reads produced by dada2 were made redundant using a custom Python script ([https://github.com/bowmanlab/seq_data_scripts](https://github.com/bowmanlab/seq_data_scripts)) and used as input for the paprica pipeline for microbial community structure and metabolic inference (Bowman and Ducklow, 2015). The paprica method for determining microbial community structure differs from most methods in that it relies on the placement of reads on a phylogenetic tree created from the 16S rRNA gene reads from all completed bacterial and archaeal genomes in Genbank. Because the metabolic potential of each phylogenetic edge on the reference tree is known, a reasonable estimate of gene content can be made for the organism of origin for each read. Edge abundance data were normalized to predict 16S rRNA gene copy number prior to statistical analysis.

Phylogenetic placement of some phylotypes of interest yielded relatively low posterior probabilities. The identity of these phylotypes was further investigated with the Ribosomal Database Project Bayesian classifier (Wang et al., 2007) and blastn megablast (Altschul et al., 1990) against the NCBI 16S rRNA gene database.

**Statistical analysis**

The count matrix generated from the 16S rRNA gene read data was used in combination with phyloxml-format phylogenetic trees generated by Guppy (Matsen et al., 2011), as implemented by paprica, along with the Archaeopteryx visualization software tool (Han and Zmasek, 2009), to identify the phylogeny of observed microbial taxa. All statistical analysis was carried out in R (R Core Team, 2012) and RStudio (RStudio Team, 2016). The R package vegan (Oksanen et al., 2013) was used to examine microbial community ecology by generating a dendrogram from the count matrix via dissimilarity analysis with hierarchical clustering based on Bray-Curtis dissimilarity index. The R package clustsig and the similarity profile analysis (SIMPROF) tool were used to determine the number of significant sample clusters based on gene content using the ‘hclust’ function, assuming no predetermined clustering (Yoshioka, 2008; Clarke et al., 2008). These results were utilized in conjunction with the gplots package and its ‘heatmap.2’ function to generate all abundance heatmaps (Warnes et al., 2016).

Principal coordinate analysis (PCoA) was performed using the ‘pcoa’ function from the R package ape v5.2
Webb et al: Impacts of Zostera eelgrasses on microbial community structure in San Diego coastal waters

Results

Bacterial and archaeal clades and potential pathogenicity

We determined significant differences in the microbial community structure across sample location and seagrass presence or absence through use of 16s rRNA sequencing, flow cytometry, and statistical analysis. SIMPROF clustering identified nine significant sample clusters from our collection of samples (n = 45), each cluster having similar overall bacterial community composition among its members (Figures 2 and 3).

The taxa identified using the ‘results’ function in DESeq2 as having a p-value < 0.05 were determined to be differentially present between the inner and outer bay samples and between eelgrass-present and absent sites. From our list of 416 identified bacterial clades, 127 of these clades were identified as differentially present between the inner and outer bay sites (Figure 2). From these 127 clades, we focused on the top 30 taxa ranked by abundance (Table 1), which enabled differences in presence and abundance of specific bacterial clades between sampling conditions to be more apparent visually (Figure 3). In outer bay samples, bacterial taxa found in significantly different abundance from inner bay samples include Fluvicola taffensis, Pseudohongiella spirulinae KCTC 32221, Sulfitobacter, Marinobacter, Pelagibacter ubique HTCC 1062, family Firmicimonadaceae, Formosa sp.Hel3 A148, Cyanobacteria, Thioglobus singularis PS1, and Planktomarina temperata RCA23 were (Table 1). In comparison, bacterial taxa found in significantly different abundance in samples from the inner bay include Martellella endophytica YC6887, Polaribacter vadi, genus Oxynema, Pelagibacter sp. IMCC9063, and order Rickettsiales (Figure 2 and Table 1).

For both the inner and outer bay, there were minor but significant differences in taxa identified between paired sites. Using DESeq2, we identified 13 differentially present bacterial clades between the eelgrass-present and eelgrass-absent sites (p < 0.05) (Table 2). Taxa found in greater abundance in samples collected at eelgrass-present sites include family Rhodobacteraceae, Candidatus Puniceispirillum marinus IMCC1322, Halologobus pacificus RR3-57, Teredinibacter turnae K7901, genus Colwellia, Thioglobus singularis PS1, genus Tenacibaculum, Fluvicola taffensis, Synechococcus sp. CC9311, and Phylum Cyanobacteria (Table 2). In comparison, bacterial taxa found in greater abundance at eelgrass-absent sites include Halobacteriovorax marinus Sj, Oscillatoria acuminata PCC 6304, and Candidatus Promineofilum breve Cfx-K (Table 2).

To further examine and explain variation in the microbial community structure of each sample (n = 45) due to the influence of eelgrass presence and sampling location, we conducted a principal coordinate analysis (PCoA). The first principal coordinate (PC1) accounted for 80.7% of the variance between samples, and the second principal coordinate (PC2) accounted for 10.1% of variance (Figure 4). The distribution of samples in PC1 is strongly dependent on sample location (inner vs. outer bay) (Student’s t-test, p approaches 0, n = 45). We therefore interpret PC1 as a good indicator of sample location. Sample distribution along PC1 is largely the result of select bacterial taxa, including Vibrio gazogenes ATCC 43942, Sulfuricurvum kuijense DSM 16994, and genus Bartonella, being elevated in the inner bay samples with positive PC1 values. Sample distribution along PC2 is largely the result of select bacterial taxa, including Altererythrobacter ishigakiensis (Paradis and Schliep, 2018), and taxon- (as phylogenetic edges or clades) weighted means for each principal coordinate were calculated with the ‘wascores’ function in package vegan (Oksanen et al., 2018). We used the R package DESeq2 (Love et al., 2014) to identify phylogenetic edges or clades that were differentially present between eelgrass-present and eelgrass-absent sites, and between sites inside and outside of San Diego Bay. DESeq2 performs differential abundance analysis based on the negative binomial/Gamma-Poisson distribution. We applied the default analysis, which uses estimation of size factors with the “median ratio method” described in Anders and Huber (2010), followed by estimation of dispersion. Next, the Wald test for generalized linear model coefficients was used to test for significance of coefficients, taking into account the size factors and dispersion estimates that were previously calculated. The DESeq2 ‘results’ (‘res’) function, with a p-value set to 0.05, allowed for testing of the null hypothesis: sampling site has no effect on which microbial taxa are present based on sampling location or eelgrass presence. This process was repeated for the domain Archaea. After the most abundant, differentially present bacterial taxa were identified, we determined which of these taxa had potential pathogens in their lower-order clades. To determine if specific clades contained bacterial strains of pathogenic bacteria, we utilized the National Institutes of Health (NIH)- National Center for Biotechnology Information (NCBI) Pathogen Detection Isolates Browser, which includes bacterial pathogen genomic sequences sourced from food, hospital patients, and the environment, as well as a literature search. Clades were labeled as potentially pathogenic if the strain itself or members of the taxonomic clade have been previously identified as being pathogenic to animals or humans.

The HNA content, LNA content, and total bacterial abundance as determined by flow cytometry were used to test for a correlation between eelgrass presence/absence and marine bacterial abundance in our samples. A mixed-effects modeling approach was applied to evaluate differences in HNA, LNA, and total bacterial abundance between sites. The lme4 package (Bates et al., 2012) was used to construct linear mixed effects models to determine and understand the interaction between eelgrass presence, bacterial abundance, and HNA and LNA content in the collected samples. For the mixed effects models, the fixed effect was “eelgrass” (present vs. absent). The random effect was an intercept for sample site location, in conjunction with by-location random slopes for the effect of eelgrass presence. Random slope models were generated to observe how the effect of eelgrass presence differed between locations (i.e., inside bay or outside bay). Significance was determined by p-values generated with likelihood ratio tests (ANOVA) against the full mixed effects models.
Of the 63 archaeal clades identified across the 45 samples, we identified five as differentially present between inner and outer bay sampling sites (Table 3). None of the differentially present archaeal clades have known pathogenic members (Table 3).

We then identified five archaea as differentially abundant between eelgrass-present vs. absent sites. As for the archaea that were differentially abundant between inner and outer bay sites, none of these are known pathogens (Table 4).

We then utilized flow cytometry data and determined whether there was a significant (p < 0.05) correlation between eelgrass presence and marine bacterial abundance in our samples. Specifically, we conducted linear regression models with random intercepts and random slopes to examine whether there was a significant correlation between eelgrass presence and bacterial abundance, with random intercepts for the effect of sample site location in conjunction with by-location random slopes for the effect of eelgrass presence. When comparing the random intercept-random slope model to the null model for HNA in a likelihood ratio test (ANOVA), we did not find evidence that eelgrass presence affected HNA abundance ($\chi^2(1) = 2.42, p = 0.1198, n = 45$) between seagrass-present and -absent sites. Similarly, when comparing the random intercept-random slope model to the null model for LNA with an ANOVA, we found no significant effect of eelgrass presence on LNA abundance ($\chi^2(1) = 2.81, p = 0.09393, n = 45$) between seagrass-present and -absent sites. In terms of overall bacterial abundance, when comparing the random intercept-random slope abundance model to the null model with an ANOVA, we could not conclude...
Figure 3: Heatmap of the most abundant, differentially present bacterial clades between inner and outer bay sites. Taxon abundance for the 30 most abundant, differentially present bacterial clades between sites is shown on a log_{10} scale for 45 samples. Darker blue indicates greater relative abundance of a specific taxon. Each color in the dendrogram represents a statistically significant cluster of samples with similar microbial community structure. Colors below the dendrogram indicate location and eelgrass presence for each sample collected: blue for inner bay, no eelgrass; pink for inner bay, eelgrass-present; green for outer bay, no eelgrass; yellow for outer bay, eelgrass present. DOI: https://doi.org/10.1525/elementa.350.f3

that eelgrass presence significantly affected overall bacterial abundance between eelgrass-present vs. -absent sites (\(\chi^2(1) = 2.91, p = 0.08799, n = 45\)).

Discussion

In this study we utilized 16S rRNA gene sequencing to examine the microbial community structure of sampling areas located within and just outside of the San Diego Bay, at sites where eelgrass beds were either present or absent. For these samples we determined that while eelgrass presence does have some influence on the microbial community structure of surrounding waters, site location (inside vs. outside of San Diego Bay) exerts a stronger influence on overall microbial community structure. These findings offer an examination of how Zostera eelgrass beds impact their ecosystem within and just outside of San Diego Bay. Regardless of eelgrass presence or absence, however, orders that may include pathogenic bacteria were relatively more abundant in samples taken from inside than from outside of the bay. This trend may result from the more quiescent conditions inside the bay, or from a greater concentration of organic carbon or nutrients, for example, from runoff or sewage input to the bay (not addressed in this study).

Among the 30 most abundant and differentially present bacterial clades between the inner and outer Bay sites, order Rickettsiales (Paddock et al., 2003; Nicholson et al., 2010; Kang et al., 2014), family Flavobacteriaceae (Touchon et al., 2011; McBride, 2014; Småge et al., 2016), and order Pseudomonadales were identified as lineages containing possible pathogens (Figure 3 and Table 1). Family Flavobacteriaceae (map id = 0.87, post-prob = 0.72) was found somewhat evenly in samples collected either within or outside of the bay as well as in eelgrass-present or -absent sites (Figure 3). Pathogens of this taxonomic family include fish pathogens Flavobacterium branchiophilum and Flavobacterium columnare and human pathogens such as Elizabethkingia meningoseptica (McBride, 2014). A phylotype classified as order Pseudomonadales (map id = 0.86, post-prob = 1.00) was evenly distributed across all sites, suggesting that eelgrass does not influence the presence of this phylotype at local scales. Some pathogenic phylotypes of this order, as determined with the NCBI Pathogen Detection Isolates Browser, include 14 strains of Acinetobacter baumannii (PHEA-2, BJAB0868, AB31, 6411, B8300, XH856, AP_882, IEC338BC, YMC2010/8/T346, CA16, SSA3, SSA6, USA15, and HUMV-6483).

Order Rickettsiales was found predominantly in samples collected from within San Diego Bay, without regard to eelgrass presence or absence. While phylo- genetic placement could not identify a specific strain, this order is of interest as it includes both human and animal pathogens such as Ehrlichia canis, Ehrlichia chafeensis, Ehrlichia ewingii, Rickettsia typhi, and Rickettsia holling iangensis (Paddock et al., 2003; Nicholson et al., 2010; Kang et al., 2014). Although it is unlikely that the recovered 16S rRNA gene reads reflect the presence of these terrestrial, tick-based pathogens, these reads could indicate the presence of alternate parasites or symbionts with a similar phylogenetic signal. This interpretation is supported by the low posterior probabilities obtained for these reads during phylogenetic placement. Members of the ubiquitous SAR11 clade of...
marine bacteria are often commonly misidentified as Rickettsiales in 16S rRNA gene studies due to similar GC bias (Luo, 2014). While our analysis did associate reads with the sequenced genomes of the SAR11 genus *Pelagibacter*, it is possible that these Rickettsiales represent a coastal clade with weak 16S rRNA gene similarity to the sequenced *Pelagibacter* genomes. Further analysis of Rickettsiales reads by blastn megablast suggested close sequence homology with members of the genus *Ehrlichia* (E-value of $2 \times 10^{-28}$, 76% identity across the alignment region); however, this analysis does not preclude misclassification.

**Table 1: Most abundant bacterial clades between inner and outer San Diego Bay sites.** DOI: https://doi.org/10.1525/elementa.350.t1

| Name                        | P-valuea | Map IDb | Posterior probabilityc | Potential pathogen | Direction of enrichmentd |
|-----------------------------|----------|---------|------------------------|--------------------|-------------------------|
| Order Rickettsiales         | 0.000    | 0.71    | 0.11                   | Yes                | Inner bay               |
| Genus Sulfitobacter         | 0.000    | 0.94    | 0.69                   | No                 | Outer bay               |
| *Planktomarina Temperata* RCA23 | $4.122 \times 10^{-36}$ | 0.96 | 0.66                   | No                 | Outer bay               |
| Order Rickettsiales         | 0.138    | 0.74    | 0.21                   | Yes                | Inner bay               |
| *Pelagibacter* sp. IMCC9063 | $2.383 \times 10^{-155}$ | 0.86 | 0.80                   | No                 | Inner bay               |
| Genus Oxynema               | 0.104    | 0.82    | 1.00                   | No                 | Inner bay               |
| *Thioglobus singularis* PS1 | $5.044 \times 10^{-91}$ | 0.95 | 0.98                   | No                 | Outer bay               |
| Family Aquifaciaceae        | $2.586 \times 10^{-52}$ | 0.64 | 0.88                   | No                 | Inner bay               |
| Genus Marinobacter          | 0.145    | 0.91    | 1.00                   | No                 | Outer bay               |
| *Fluviicola taffensis*      | $1.242 \times 10^{-63}$ | 0.85 | 0.99                   | No                 | Outer bay               |
| Phylum Planctomyces         | 0.893    | 0.90    | 0.84                   | No                 | Inner bay               |
| *Oscillaria acuminata* PCC 6304 | $1.302 \times 10^{-51}$ | 0.82 | 0.33                   | No                 | Outer bay               |
| *Wenzhouxiangella* marina KCTC 42284 | $1.806 \times 10^{-48}$ | 0.88 | 0.64                   | No                 | Outer bay               |
| *Pseudohongiella* spirulinae KCTC 32221 | $1.081 \times 10^{-44}$ | 0.90 | 0.99                   | No                 | Outer bay               |
| *Polaribacter* vadi         | $3.070 \times 10^{-35}$ | 0.87 | 0.61                   | No                 | Inner bay               |
| *Formosa* sp. He13 A1 48    | $1.357 \times 10^{-33}$ | 0.86 | 0.98                   | No                 | Outer bay               |
| *Maritellella* endophytica* YC6887 | $3.955 \times 10^{-28}$ | 0.94 | 0.45                   | No                 | Inner bay               |
| Phylum Cyanobacteria        | $3.287 \times 10^{-23}$ | 0.78 | 0.66                   | No                 | Outer bay               |
| Genus Acetobacter           | $1.038 \times 10^{-22}$ | 0.48 | 0.89                   | No                 | Outer bay               |
| Alpha Proteobacterium* HIMB59 | $9.553 \times 10^{-20}$ | 0.86 | 0.94                   | No                 | Inner bay               |
| *Synechocystis* sp.         | $1.841 \times 10^{-39}$ | 0.85 | 0.36                   | No                 | Outer bay               |
| Candidatus Pelagibacter* ubique HTCC1062 | $1.060 \times 10^{-38}$ | 0.93 | 0.93                   | No                 | Outer bay               |
| Genus Pelagibacter          | $2.432 \times 10^{-38}$ | 0.84 | 0.38                   | No                 | Inner bay               |
| *Microcystis* aeruginosa    | $4.855 \times 10^{-35}$ | 0.81 | 0.77                   | No                  | Inner bay               |
| Family Rhodobacteraceae     | $4.496 \times 10^{-30}$ | 0.95 | 0.66                   | No                 | Outer bay               |
| Order Pseudomonadales       | $1.933 \times 10^{-9}$  | 0.86 | 1.00                   | Yes                | Inner bay               |
| Family Fimbriomonadales     | $1.764 \times 10^{-7}$  | 0.78 | 0.06                   | No                 | Outer bay               |
| Phylum Actinobacteria       | $1.655 \times 10^{-6}$  | 0.80 | 0.61                   | No                 | Inner bay               |
| *Halloglobus* pacificus RR3-57 | $9.535 \times 10^{-5}$  | 0.88 | 0.53                   | No                 | Outer bay               |
| Family Flavobacteriaceae    | $7.700 \times 10^{-3}$  | 0.87 | 0.72                   | Yes                | Inner bay               |

*a* P-value $< 0.05$ indicates strong evidence against the null hypothesis that sampling location has no influence on the microbial community structure at each of the collection sites.

*b* Map ID indicates the fraction of nucleotide bases in the 16S rRNA targeted amplicon reads that matched the reference clade (average for each clade across all samples).

*c* Posterior probability indicates the statistical likelihood that the phylogenetic placement is correct.

*d* Direction of enrichment indicates where each clade or taxon was found in greater average abundance: inner bay vs. outer bay.

*e* Known to form harmful algal blooms (Jacoby et al., 2000).
Table 2: Differentially present bacterial clades between seagrass-present vs. absent San Diego Bay sites. DOI: https://doi.org/10.1525/elementa.350.t2

| Name                                      | P-Value | Map ID | Posterior probability | Potential pathogen | Direction of enrichment |
|-------------------------------------------|---------|--------|-----------------------|--------------------|-------------------------|
| Family Rhodobacteraceae                   | 0.0044  | 0.95   | 0.66                  | No                 | EG+                     |
| Candidatus Puniceispirillum marinum IMCC1322 | 0.0176  | 0.90   | 1.00                  | No                 | EG+                     |
| Halioglobus pacificus RR3-57              | 0.0481  | 0.88   | 0.53                  | No                 | EG+                     |
| Teredinibacter turneri T7901              | 0.0110  | 0.92   | 0.96                  | No                 | EG+                     |
| Genus Colwellia                           | 0.0212  | 0.95   | 0.86                  | No                 | EG+                     |
| Thioglobus singularis PS1                 | 0.0241  | 0.95   | 0.98                  | No                 | EG+                     |
| Halobacteriovorax marinus Sj              | 0.0255  | 0.81   | 0.92                  | No                 | EG–                     |
| Genus Tenacibaculum                       | 0.0303  | 0.99   | 0.60                  | Yes                | EG+                     |
| Fluvicola taaffensis                      | 0.0258  | 0.85   | 0.99                  | No                 | EG+                     |
| Synechococcus sp. CC9311                  | 0.0345  | 0.98   | 0.92                  | No                 | EG+                     |
| Phylum Cyanobacteria                      | 0.0248  | 0.77   | 0.65                  | No                 | EG+                     |
| Oscillatoria acuminata PCC 6.304          | 0.0139  | 0.82   | 0.33                  | No                 | EG–                     |
| Candidatus Promineofilum breve Cfx-K      | 0.0435  | 0.78   | 0.35                  | No                 | EG–                     |

- P-value < 0.05 indicates strong evidence against the null hypothesis that eelgrass presence has no influence on the microbial community structure at each of the collection sites.
- Map ID indicates the fraction of nucleotide bases in the 16S rRNA targeted amplicon reads that matched the reference clade (average for each clade across all samples).
- Posterior probability indicates the statistical likelihood that the phylogenetic placement is correct.
- Direction of enrichment indicates where each clade or taxon was found in greater average abundance: sampling site with eelgrass present (EG+) or eelgrass absent (EG–).

Figure 4: Principal coordinate analysis plot for all bacterial clades. Distribution of the samples in a space defined by the first two principal coordinates (PC1 and PC2). PC1 accounted for 80.7% of variance, and the samples are distributed in PC1 largely according to sampling location. PC2 accounted for a 10.1% proportion of variance between samples. DOI: https://doi.org/10.1525/elementa.350.f4
In contrast to Lamb et al. (2017), we did not find conclusive evidence that eelgrass beds reduced overall microbial contamination in San Diego coastal waters. This discrepancy could reflect geographical differences; the Indonesian locations where Lamb et al. (2017) conducted their research are exposed to significant levels of microbial contamination of the San Diego coastal environment. Sewage contamination of the Indonesian islands could also account for differences, with a “ring” of eelgrass beds around the islands reducing the transport of pathogens outside the ring. Sewage contamination of the San Diego coastal environment is associated with significant rain events, with the last heavy rainfall prior to sampling coming nearly 6 months earlier on 9 January 2017 (https://www.ncdc.noaa.gov/stormevents). The difference in pathogen load between eelgrass-present and eelgrass-absent sites may be much more pronounced immediately following rain events.

Regardless of any overall reduced pathogen load, we did identify 13 bacterial taxa that were differentially present between eelgrass-present and eelgrass-absent sites. While none of these taxa are known to include human pathogens, some are of interest for other reasons. *Teredinibacter turnerae* is known to exhibit antibiotic activity against both Gram-positive and Gram-negative bacteria including *Bacillus cereus* and *Staphylococcus sciuri* (Trindade-Silva et al., 2009). *Halobacteriovorax marinus* SJ and other members of genus *Halobacteriovorax* are predatory bacteria that attack Gram-negative bacteria including *Vibrio*, *E. coli*, and *Pseudomonas* variety (Enos et al., 2018). *Genus Tenacibaculum* contains a pathogenic phylotype, *Tenacibaculum maritimum*, which is a known fish pathogen that infects fish kidneys and causes symptoms including skin lesions and fin rot (Avendaño-Herrera et al., 2006). *Teredinibacter turnerae* and *Tenacibaculum maritimum* were significantly more abundant in samples collected from eelgrass-present sites, while *Halobacteriovorax marinus* SJ was significantly more abundant at eelgrass-absent sites. Each of these phylogenetic placements had relatively high posterior probabilities (0.96, 0.60, and 0.92 respectively), indicating high confidence in the phylogenetic placement. We note that bacteria that are potentially beneficial to human health were present at both eelgrass-present

---

**Table 3:** Differentially present archaeal clades between inner and outer San Diego Bay sites. DOI: https://doi.org/10.1525/elementa.350.t3

| Name                          | P-value<sup>a</sup> | Map ID<sup>b</sup> | Posterior probability<sup>c</sup> | Potential pathogen | Direction of enrichment<sup>d</sup> |
|-------------------------------|---------------------|--------------------|-----------------------------------|-------------------|-------------------------------------|
| Domain Archaea                | 0.0006              | 0.150              | 0.676                             | No                | Outer bay                           |
| Order Methanococcales         | 0.002               | 0.626              | 0.075                             | No                | Outer bay                           |
| *Methanospirillum hungatei* JF-1 | 0.037              | 0.642              | 0.074                             | No                | Outer bay                           |
| *Methanococcus maripaludis* S2 | 0.046               | 0.645              | 0.106                             | No                | Outer bay                           |
| Kingdom Proteoarchaeota-TACK  | 0.047               | 0.665              | 0.076                             | No                | Outer bay                           |

<sup>a</sup> P-value < 0.05 indicates strong evidence against the null hypothesis that sampling location has no influence on the microbial community structure at each of the collection sites.

<sup>b</sup> Map ID indicates the fraction of nucleotide bases in the 16S rRNA targeted amplicon reads that matched the reference clade (average for each clade across all samples).

<sup>c</sup> Posterior probability indicates the statistical likelihood that the phylogenetic placement is correct.

<sup>d</sup> Direction of enrichment indicates where each clade or taxon was found in greater average abundance: inner bay vs. outer bay.

---

**Table 4:** Differentially present archaeal clades between seagrass-present vs. -absent San Diego Bay sites. DOI: https://doi.org/10.1525/elementa.350.t4

| Name                          | P-value<sup>a</sup> | Map ID<sup>b</sup> | Posterior probability<sup>c</sup> | Potential pathogen | Direction of enrichment<sup>d</sup> |
|-------------------------------|---------------------|--------------------|-----------------------------------|-------------------|-------------------------------------|
| Domain Archaea                | 0.0006              | 0.150              | 0.676                             | No                | SG--                                |
| Order Methanococcales         | 0.002               | 0.626              | 0.075                             | No                | SG--                                |
| *Methanospirillum hungatei* JF-1 | 0.037              | 0.642              | 0.074                             | No                | SG+                                 |
| Genus Ferroplasma             | 0.046               | 0.585              | 0.470                             | No                | SG--                                |
| Kingdom Proteoarchaeota-TACK  | 0.047               | 0.665              | 0.076                             | No                | SG--                                |

<sup>a</sup> P-value < 0.05 indicates strong evidence against the null hypothesis that eelgrass presence has no influence on the microbial community structure at each of the collection sites.

<sup>b</sup> Map ID indicates the fraction of nucleotide bases in the 16S rRNA targeted amplicon reads that matched the reference clade (average for each clade across all samples).

<sup>c</sup> Posterior probability indicates the statistical likelihood that the phylogenetic placement is correct.

<sup>d</sup> Direction of enrichment indicates where each clade or taxon was found in greater average abundance: sampling site with eelgrass present (EG+) or eelgrass absent (EG–).
and absent sites; i.e., the presence of Teredinibacter turnerae at sites with eelgrass and Halobacteriovorax marinus SJ at sites without. Due to their antibiotic and predatory nature against human pathogens, further research into the presence of these bacteria would allow an examination of how they may be altering the local marine microbiome, whether in conjunction with Zostera eelgrass beds or not, and how their characteristics might be leveraged for biotechnological applications.

Eelgrass beds are critical components of the coastal marine ecosystem, with demonstrated influence on microbial community structure. In some environments, however, their influence on microbial community structure may be less significant than other factors, e.g., site location, physical processes, or other external factors not addressed in this experiment. Although we are cautious about extrapolating our findings to other coastal environments, this experiment presents useful information regarding specific microbial community structure of San Diego Bay and its immediately adjacent waters. Follow-on work to identify differences in microbial community structure when pathogen load is elevated, such as after heavy rainfall, and across an expanded selection of sites, will further clarify the influence on eelgrass beds on microbial community structure.

Data Accessibility Statement
DNA sequences: NCBI SRA: PRJNA483963.

Supplemental file
The supplemental file for this article can be found as follows:

- Table S1. Sample Data. DOI: https://doi.org/10.1525/elementa.350.s1

Acknowledgements
We would like to thank Brett Pickering for assistance with sampling and Jane Teranes for facilitating the SURF REU program.

Funding information
TR was supported by the Scripps Institution of Oceanography SURF REU program (NSF-OCE 1659793). JSB was supported by the Simons Foundation Early Career Investigator in Marine Microbial Ecology and Evolution fellowship. Additional support came from the Doherty Career Development Fund through Scripps Institution of Oceanography.

Competing interests
The authors have no competing interests to declare.

Author contributions
- Contributed to conception and design: <TR, JSB>
- Contributed to acquisition of data: <CTR, JSB, NE>
- Contributed to analysis and interpretation of data: <SJW, JSB>
- Drafted and/or revised the article: <SJW, JSB>
- Approved the submitted version for publication: <SJW, TR, NE, JSB>

References
Alam, K, Agua, T, Maven, H, Taie, R, Rao, KS, Burrows, I, Huber, ME and Rali, T. 1994. Preliminary screening of seaweeds, seagrass and lemongrass oil from Papua New Guinea for antifungal and antiviral activity. Int J Pharmocogn Phytochem Res 3(2): 396–399. DOI: https://doi.org/10.3109/1388029409083022
Altschul, SF, Gish, W, Miller, W, Myers, EW and Lipman, DJ. 1990. Basic local alignment search tool. J Mol Biol 215(3): 403–410. DOI: https://doi.org/10.1016/S0022-2836(05)80360-2
Avendaño-Herrera, R, Toranzo, AE and Magariños, B. 2006. Tenacibaculosis infection in marine fish caused by Tenacibaculum maritimum: A review. Dis Aquat Organ 71(3): 255–266 DOI: https://doi.org/10.3354/dao071255
Bates, DM, Maechler, M and Bolker, B. 2012. lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-0.
Bowman, JS, Amaral-Zettler, L, Rich, J, Luria, C and Ducklow, H. 2017. Segmentation of the bacterial community facilitates the prediction of ecosystem function along the western Antarctic Peninsula. ISME J 11(6): 1460–1471. DOI: https://doi.org/10.1038/ismej.2016.204
Bowman, JS and Ducklow, HW. 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. DOI: https://doi.org/10.1371/journal.pone.0135868
Callahan, BJ, McMurdie, PJ, Rosen, MJ, Han, AW, Johnson, AJA and Holmes, SP. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods 7(7): 581–583. DOI: https://doi.org/10.1038/nmeth.3869
Choi, HG, Lee, JH, Park, HH and Sayegh, FAQ. 2009. Antioxidant and antimicrobial activity of Zostera marina L. extract. Algae 24(3): 179–184. DOI: https://doi.org/10.4490/ALGAE.2009.24.3.179
Clarke, KR, Somerfield, PJ and Gorley, RN. 2008. Testing of null hypotheses in exploratory community analyses similarities profiles and biota-environment linkage. J Exp Mar Biol Ecol 366(1–2): 56–69. DOI: https://doi.org/10.1016/j.jembe.2008.07.009
Enos, BG, Anthony, MK, DeGiorgis, JA and Williams, LE. 2018. Prey range and genome evolution of Halobacteriovorax marinus predatory bacteria from an estuary. MSphere 3(1): e00508–17. DOI: https://doi.org/10.1128/mSphere.00508-17
Eren, AM, Vineis, JH, Morrison, HG and Sogin, ML. 2013. A filtering method to generate high quality short reads using illumina paired-end technology. PLoS One 8(6): e66643. DOI: https://doi.org/10.1371/journal.pone.0066643
Ettinger, CL, Voerman, SE, Lang, JM, Stachowicz, JJ and Eisen, JA. 2017. Microbial communities in sediment from Zostera marina patches, but not the Z. marina
leaf or root microbiomes, vary in relation to distance from patch edge. PeerJ 5: e3246. DOI: https://doi.org/10.7717/peerj.3246

Fairey, R, Roberts, C, Jacobi, M, Lamerdin, S, Clark, R, Downing, J, Long, E, Hunt, J, Anderson, B, Newman, J, Tjeerdema, R, Stephenson, M and Wilson, C. 1998. Assessment of sediment toxicity and chemical concentrations in the San Diego Bay region, California, USA. Environ Toxicol Chem/SETAC 17(8): 1570–1581. DOI: https://doi.org/10.1002/etc.5620170819

Han, MV and Zmasek, CM. 2009. phyloXML: XML for evolutionary biology and comparative genomics. BMC Bioinformatics 10: 356. DOI: https://doi.org/10.1186/1471-2105-10-356

Hogarth, PJ. 2015. The Biology of Mangroves and Seagrasses. Oxford. DOI: https://doi.org/10.1093/acpr/9780198716549.001.0001

Kang, YJ, Diao, XX, Zhao, GY, Chen, MH, Xiong, Y, Shi, M, Fu, WM, Guo, YJ, Pan, B, Chen, XP, Holmes, EC, Gillespie, JJ, Dumlier, SJ and Zhang, YZ. 2014. Extensive diversity of Rickettsiales bacteria in two species of ticks from China and the evolution of the Rickettsiales. BMC Evol Biol 14(1): 1–12. DOI: https://doi.org/10.1186/s12862-014-0167-2

Kannan, RRR, Arumugam, R and Anantharaman, P. 2010. Antibacterial potential of three seagrasses against human pathogens. Asian Pac J Trop Med 3(11): 890–893. DOI: https://doi.org/10.1016/S1995-7645(10)60214-3

Lamb, JB, van de Water, JAJM, Bourne, DG, Altier, C, Hein, MY, Fiorenza, EA, Abu, N, Jompa, J and Harvell, CD. 2017. Seagrass ecosystems reduce exposure to bacterial pathogens of humans, fishes, and invertebrates. Science 355(6326): 731–733. DOI: https://doi.org/10.1126/science.aal1956

Love, MI, Huber, W and Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15: 550. DOI: https://doi.org/10.1186/s13059-014-0550-8

Luo, H. 2014. Evolutionary origin of a streamlined marine bacterioplankton lineage. ISME J 9(6): 1423–1433. DOI: https://doi.org/10.1038/ismej.2014.227

Mayavu, P, Sugesh, S and Ravindran, VJ. 2009. Antibacterial activity of seagrass species against biofilm forming bacteria. Res J Microbiol 4(8): 314–319. DOI: https://doi.org/10.3923/jm.2009.314.319

McBride, MJ. 2014. The Family Flavobacteriaceae. In: Rosenberg, E, DeLong, EF, Lory, S, Stackebrandt, E and Thompson, F (eds.), The Prokaryotes. 4th ed. Berlin, Heidelberg: Springer. DOI: https://doi.org/10.1007/978-3-642-38954-2_130

Merkel & Associates, Inc. 2014. 2014 San Diego Bay Eelgrass Inventory Update. [Online] Available at: https://www.westcoast.fisheries.noaa.gov/publications/gis_maps/maps/eelgrass/2014-eelgrass-survey-report.pdf. [Accessed 15 January 2018]

Nicholson, WL, Allen, KE, McQuiston, JH, Breitachwerdt, EB and Little, SE. 2010. The increasing recognition of rickettsial pathogens in dogs and people. Trends Parasitol 26(4): 205–212. DOI: https://doi.org/10.1016/j.pt.2010.01.007

NOAA-National Oceanic and Atmospheric Administration. 2018. Seagrasses. [Online] Available at: http://www.westcoast.fisheries.noaa.gov/habitat/fish_habitat/seagrass_2.html. [Accessed 15 January 2018]

Oksanen, J, Blanchet, FG, Kindt, R, Legendre, P, McGlinn, D, Minchin, PR, O’Hara, RB, Simpson, GL, Solymos, P, Stevens, MHH, Szoecs, E and Wagner, H. 2013. Vegan: Community ecology package. R package version 2.0-8. [Online] Available at: http://CRAN.R-project.org/package=vegan

Ortega, C, Solo-Gabriele, HM, Abdelzaher, A, Wright, M, Deng, Y and Stark, LM. 2009. Correlations between microbial indicators, pathogens, and environmental factors in a subtropical estuary. Mar Poll Bull 58(9): 1374–1381. DOI: https://doi.org/10.1016/j.marpolbul.2009.04.015

Paddock, CD and Childs, JE. 2003. *Ehrlichia chaffeensis*: A prototypical emerging pathogen. Clin Microbiol Rev 16(1): 37–64. DOI: https://doi.org/10.1128/CMR.16.1.37-64.2003

Paradis, E and Schliep, K. 2018. ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics. July, 1–3. DOI: https://doi.org/10.1093/bioinformatics/bty633

RStudio Team. 2016. RStudio: Integrated development for R. Boston, MA: RStudio, Inc. [Online] Available at: http://www.rstudio.com/

San Diego County. 2018. Beach and bay water quality program. [Online] Available at: http://www.sdbeachinfo.com/#SanDiegoCounty.gov. [Accessed 15 January 2018]

Småge, SB, Brevik, Ø, Duesund, H, Ottem, KF, Watanabe, K and Nylund, A. 2016. *Tenacibaculum fanmarkense* sp. nov., a fish pathogenic bacterium of the family Flavobacteriaceae isolated from Atlantic salmon. Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology 109(2): 273–285. DOI: https://doi.org/10.1007/s10482-015-0630-0

Touchon, M, Barbier, P, Bernardet, JF, Loux, V, Vacherie, B, Barbe, V, Rocha, EP and Dudaed, E. 2011. Complete genome sequence of the fish pathogen *Flavobacterium branchiophilum*. Appl Environ Microbiol 77(21): 7656–7662. DOI: https://doi.org/10.1128/AEM.05625-11

Trindade-Silva, A, Machado-Ferreira, E, Senra, M, Vizzoni, V, Yparraiguirre, L, Leoncini, O and Soares, C. 2009. Physiological traits of the symbiotic bacterium *Teredinibacter turnerae* isolated from the mangrove shipworm Neoteredo reynei. Gen Molec Biol 32(3): 572–581. DOI: https://doi.org/10.1590/S1415-475720090003000061

U.S. Department of Justice – Environment and Natural Resources Division. 2018. [Online] Letter Available at: https://therevelator.org/wp-content/uploads/2018/03/Justice-Department-Letter-March-1-IBWC.pdf. [Accessed 15 January 2018]
Walters, W, Hyde, ER, Berg-Lyons, D, Ackermann, G, Humphrey, G, Parada, A, Gilbert, J, Jansson, J, Caporaso, JG, Fuhrman, J, Apprill, A and Knight, R. 2016. Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *MSystems* 1(1). DOI: https://doi.org/10.1128/mSystems.00009-15

Wang, Q, Garrity, GM, Tiedje, JM and Cole, JR. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73(16): 5261–5267. DOI: https://doi.org/10.1128/AEM.00062-07

Warnes, GR, Bolker, B, Bonebakker, L, Gentleman, R, Liaw, WHA, Lumley, T, Maechler, M, Magnusson, A, Moeller, S, Schwartz, M and Venables, B. 2016. Package “gplots”: Various R programming tools for plotting data. *R Package Version 2.17.0* 1–68.

Yoshioka, PM. 2008. Misidentification of the Bray-Curtis similarity index. *Mar Ecol Prog Ser* 368: 309–310. DOI: https://doi.org/10.3354/meps07728