Seagrass mediates microalgal community structure at a distance

Emily Jacobs-Palmer[1], Ramón Gallego[1,2], Ana Ramón-Laca[2,3], Emily Kunselman[4], Kelly Cribari[1], Micah Horwith[4], Ryan Kelly[1].

1. University of Washington School of Marine and Environmental Affairs, Seattle, WA.
2. National Oceanic and Atmospheric Administration Northwest Fisheries Science Center, Seattle, WA.
3. Ocean Associates, Inc., Arlington, VA.
4. Washington State Department of Natural Resources, Olympia, WA.

Abstract

Seagrass beds provide a variety of ecosystem services, some of which accrue outside the bounds of the habitat itself. Here we use environmental DNA (eDNA) amplicons to analyze the temporal and spatial effect of eelgrass (*Zostera marina*) on the immediately surrounding ecological community. Sampling seawater along transects extending outward from eelgrass beds, we demonstrate that eDNA provides meter-scale resolution of communities in the field. We evaluate eDNA abundance indices for twelve major phylogenetic groups of marine and estuarine taxa along these transects, finding highly local changes linked with proximity to *Z. marina* for a diverse group of dinoflagellates. Eelgrass habitat consistently and dramatically limits dinoflagellate abundance both within the beds and for at least fifteen meters outside. Because

many dinoflagellates are capable of forming Harmful Algal Blooms (HABs) toxic to humans and other animal species, the salutory effect of eelgrass habitat on neighboring waters has important implications for public health as well as shellfish aquaculture and harvesting.

Introduction

Seagrass species throughout the world's oceans are ecosystem engineers (Jones, Lawton & Shachak, 1994), generating and sustaining habitat for a multitude of associated taxa (Duffy, 2006). Additionally, these marine macrophytes provide a wide variety of essential ecosystem services that directly benefit humans, such as temporary carbon sequestration (Fourqurean et al., 2012), nursery habitat for human food species (Heck Jr, Hays & Orth, 2003), and coastal protection through sediment accretion and stabilization (Potouroglou et al., 2017; reviewed in Nordlund et al., 2016). Some benefits, such as reduced exposure to pathogens, have been shown to accrue to organisms and ecosystems even outside of seagrass habitat boundaries (e.g. Lamb et al., 2017).

Eelgrass (*Zostera marina*) is the dominant seagrass along temperate coasts of the Northern hemisphere (Short et al., 2007). Recent worldwide declines in this species and other seagrass taxa are alarming (Orth et al., 2006; but see Shelton et al., 2017), and have been met with local protection measures in some cases, such as designation of seagrass as a 'Habitat Area of Particular Concern' (see NOAA Fisheries), as well as a Puget Sound 'Vital Sign' indicator species (Puget Sound Partnership) and the target of 'no net loss' policies (NOAA Fisheries, 2014). Frequently, a tradeoff between eelgrass conservation and aquaculture is presumed when such conservation efforts compete with shellfish seeding grounds (Hosack et al., 2006). However, commercially important species such as oysters are in fact often proximally associated

with *Z. marina* beds in the wild; they may thus depend on services provided by the habitat, and vice versa (for example, see Groner et al., 2018).

To examine how *Z. marina* modifies the biological community existing both within and immediately surrounding the habitat itself, we use environmental DNA (eDNA) from water samples to survey the presence and relative abundance of organisms on a series of alongshore transects in the coastal or estuarine waters of Washington State. Each transect extends from within eelgrass beds to bare substrate and was sampled at three timepoints during the late spring and summer. By a large margin, we find that dinoflagellates are the group most affected by *Z. marina*; spatial proximity to eelgrass habitat is associated with a taxonomically widespread decrease in dinoflagellate abundance for meters outside the borders of the beds. These results extend previous evidence for an allelopathy of *Z. marina* (and/or associated taxa) towards particular harmful algal bloom (HAB) species that cause paralytic or diarrhetic shellfish poisoning (e.g. Inaba et al., 2017), by demonstrating an effect of eelgrass communities on dinoflagellates in a 'halo' of influence surrounding the habitat. In the region of study, toxigenic dinoflagellate distributions have expanded over time, and are associated with an increase in the number of shellfish harvesting closures (Trainer et al., 2003; Moore et al., 2009). Far-reaching effects of eelgrass communities on HAB-producing taxa could therefore strengthen connections between seagrass habitat and human health, particularly in native communities with elevated rates of shellfish consumption (Washinton State Department of Ecology, 2013). Thus, our findings may carry potentially critical ramifications for management of both seagrass and shellfish in the many regions of the world where the two coincide.

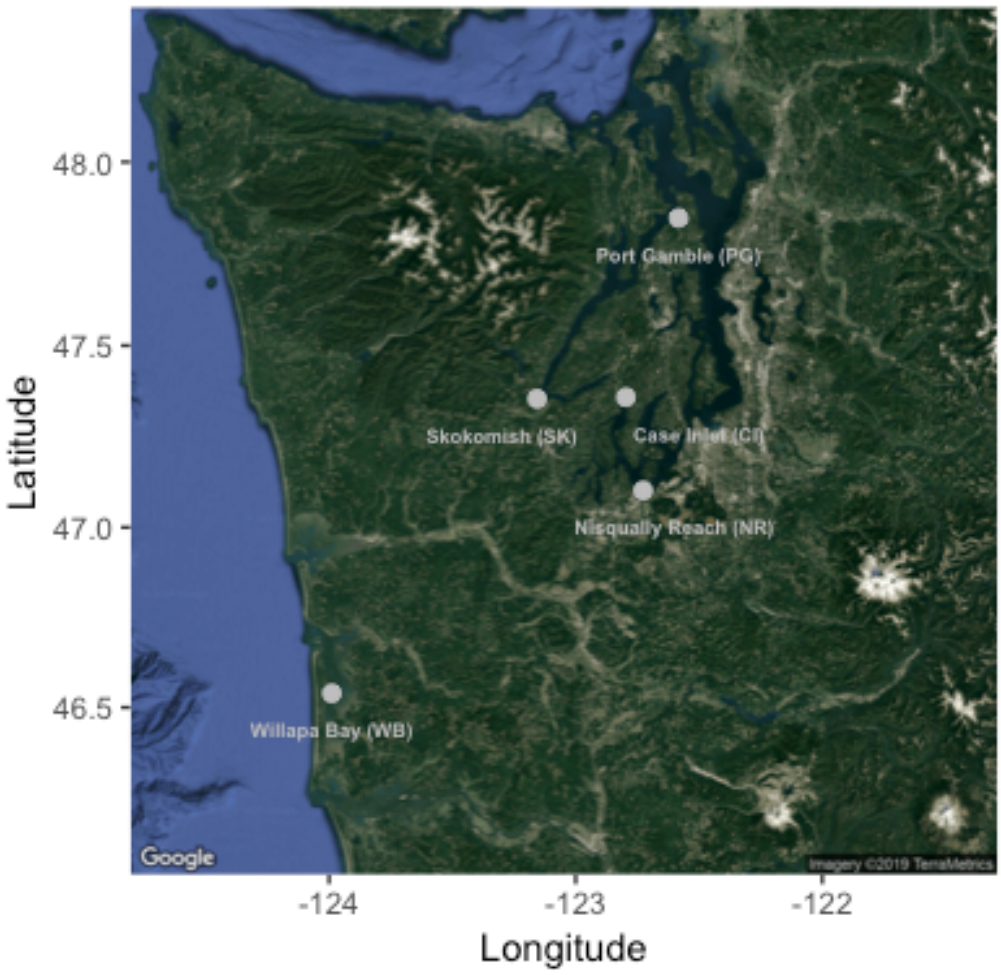
Methods

Environmental DNA sample collection

Environmental DNA sequenced with a single genetic locus provides an assay of community composition consisting of many taxa. The design of the particular PCR primers used largely determines the taxonomic composition, but it is not uncommon to sequence hundreds of taxa from dozens of phyla in a given sampling effort. Here, we targeted a ca. 313 bp fragment of COI using a primer set (Leray et al., 2013) known to amplify a broad range of marine taxa including diatoms, dinoflagellates, metazoans, fungi, and others; this primer set is broadly used in ecological applications (e.g. Leray & Knowlton, 2015; Gibson et al., 2014).

To determine the biological community composition within *Z. marina* beds and the surrounding habitat from eDNA, we sampled seawater from five sites in Puget Sound: Port Gamble, Case Inlet, Nisqually Reach, Skokomish, and Willapa Bay (Figure 1). We surveyed each location at three timepoints during the summer season, in May, July, and August of 2017. Specifically, we collected a 1 liter bottle of seawater immediately under the water surface from the approximate center of the beds ("eelgrass"), from each point in a transect extending alongshore at 1, 3, 6, 10, and 15m from the edge of the beds, and from a final location from which seagrass was absent ("bare") between 16 and 670m from the beds (the edge of each bed was defined as the point at which shoot density fell below 3 shoots/m^2; see Table S1 for precise transect locations by site). Due to local geography and conditions, it was not always possible to gather all transect samples during each sampling event; a comprehensive list of samples gathered is given in Table S1. We kept samples on ice until processing by filtering 500mL from each sample under vacuum pressure through a cellulose acetate filter with 47 mm diameter and 0.45 um pore size and stored

85 the filter at room temperature in Longmire's buffer (Renshaw et al., 2015).The final dataset
86 consisted of 84 water samples.



87

88 *Figure 1: Nearshore sampling locations in Puget Sound and outer coast, Washington, USA. GPS*
89 *coordinates are given in Supplemental Table 1.*

90 Extraction and amplification

91 To extract DNA from the sample filters, we used a phenol:chloroform:isoamyl alcohol protocol
92 (Renshaw et al., 2015), resuspended the eluate in 200 uL water, and used 1 uL of diluted DNA
93 extract (between 1:10 and 1:400) as template for PCR. To survey the eukaryotic organisms

present in our samples, PCR reactions from each of the 84 biological samples were run and sequenced in triplicate to distinguish technical from biological variance. To sequence many samples and their replicates in a single run while avoiding amplification bias due to index sequence, we followed a two-step PCR protocol (O'Donnell et al., 2016). In the first step, we used a PCR reaction containing 1X HotStar Buffer, 2.5 mM MgCl2, 0.5 mM dNTP, 0.3 µM of each primer and 0.5 units of HotStar Taq (Qiagen Corp., Valencia, CA, USA) per 20 µL reaction. The PCR protocol for this step consisted of 40 cycles, including an annealing touchdown from 62 ◦C to 46 ◦C (-1 ◦C per cycle), followed by 25 cycles at 46 ◦C. In the second step, we used a similar PCR reaction, but substituted primers with extra 5' 6-base pair tags to index samples, and a similar but shorter protocol with only 10 cycles at 46 ◦C. Finally, we generated amplicons with the same replication scheme for both positive (kangaroo (genus *Macropus*) or ostrich (genus *Struthio*) tissue, selected because these species are absent from the sampling sites, and thus we could identify cross-contamination using reads from these taxa) and negative controls (molecular grade water), and verified by gel electrophoresis that negative controls contained no appreciable amount of DNA.

Sequencing

To prepare libraries of replicated, indexed samples and positive controls, we followed manufacturers' protocols (KAPA Biosystems, Wilmington, MA, USA; NEXTflex DNA barcodes; BIOO Scientific, Austin, TX, USA). We then performed sequencing on an Illumina MiSeq (250-300 bp, paired-end) platform in four different sets of samples: two MiSeq V.2 runs and two MiSeq V.3 runs. We processed each batch separately through the initial bioinformatics analysis (see below). We employed hierarchical clustering on transects containing six PCR

replicates sequenced across two different runs (three technical replicates per run derived from the same sampled bottle of water) and found that these samples were each others' nearest neighbors (Figure S1)); thus sequencing-run-level effects were negligible and we combined the data from the four sequencing runs.

Bioinformatics

We followed updated versions of previously published procedures for bioinformatics, quality-control, and decontamination (Kelly, Gallego & Jacobs-Palmer, 2018). This protocol uses a custom Unix-based script (Gallego) calling third-party programs to perform initial Quality Control (QC) on sequence reads from all four runs combined, demultiplexing sequences to their sample of origin and clustering of unique variants into Sequence Variants (ASVs) (Martin, 2011; Callahan et al., 2016). The output is a dataset including counts of each ASV per PCR replicate; ~28M sequence reads from 19370 ASVs emerged from this step. To address possible cross-sample contamination (see Schnell, Bohmann & Gilbert, 2015), we subtracted the maximum proportional representation of each ASV across all control samples (sequenced from extraction of kangaroo or ostrich tissue) from the respective ASV in field samples; 27M reads from 19320 ASVs passed this step. After removing the two PCR replicates with an extremely low number of reads, we estimated the probability of ASV occurrence by performing site-occupancy modeling using multiple PCR replicates from each environmental sample as independent draws from a common binomial distribution, and discarded ASVs with <0.8 estimated probability; 25M reads from 3143 ASVs survived this step (Royle & Link, 2006; Lahoz-Monfort, Guillera-Arroita & Tingley, 2015). Lastly, we removed samples whose PCR replicates were highly dissimilar: we calculated the Bray-Curtis dissimilarity amongst PCR replicates from the same bottle of water

and discarded those with distance to the sample centroid outside a 95% Confidence Interval. Of 84 bottles of water collected, 3 technical replicates survived QC in 72 cases (86%), 2 replicates in 9 cases (11%), 1 replicate in 2 cases (2%), and zero replicates in a single case (1%) (Table S1). The final dataset of 24M reads from 3142 ASVs comprised 83% of the original sequence reads.

All bioinformatic and analytical code is included in this manuscript, and provides the details of parameter settings in the bioinformatics pipelines used. Sequence and annotation information are included as well, and the former are deposited and publicly available in GenBank (upon acceptance).

Taxonomy

To assign taxonomy to each ASV sequence, we followed the protocol detailed in Kelly, Gallego & Jacobs-Palmer (2018). Briefly, this protocol uses blastn (Camacho et al., 2009) on a local version of the full NCBI nucleotide database (current as of February 13, 2019), recovering up to 100 hits per query sequence with at least 85% similarity and maximum e-values of 10^{-30} (culling limit = 5), and reconciling conflicts among matches using the last common ancestor approach implemented in MEGAN 6.4 (Huson et al., 2016). Within MEGAN, we imposed an additional more stringent round of quality-control to ensure sufficient similarity between query and database sequences by requiring a bit score of at least 450 (ca. 90% identical over the entire 313bp fragment). Of the 24M total reads in our dataset, we were able to annotate 4.1m to the level of phylum or lower; the majority of the remaining reads had no BLAST hits meeting our criteria (7.6M) or else did not receive taxonomic assignment due to insufficient similarity or

159 conflicting BLAST hits (12.1M). We use the annotated sequences in our taxonomic analyses
160 below .

161 Because dinoflagellates had different ecological patterns than other taxa (see Results), we further
162 refined our annotations for these ASVs. For sequence variants both a) assigned to a taxon within
163 Dinoflagellata, and b) having more than a trivial number of reads in the dataset (> 1000), we
164 considered the geographic range of taxa involved (restricting possible annotations to those taxa
165 known from the North Pacific) and assigning taxonomy conservatively only in cases of >97%
166 sequence identity between the subject and query sequence. Three distinct dinoflagellate
167 sequences with identical amino-acid translations from the genus *Heterocapsa* co-occurred in
168 time and space; to avoid pseudoreplication, we treated these as a single taxonomic unit (this
169 choice did not affect the trends or significance of results). A phylogeny built of the eleven
170 remaining dinoflagellate sequences (Figure S4) confirmed that family- and genus-level
171 taxonomic groups occupied monophyletic clades (Li et al., 2015).

172 Statistical Analysis

173 Community Composition

174 To confirm the spatial resolution of our eDNA communities, we used non-metric
175 multidimensional scaling (nMDS) ordination of eDNA indices for all ASVs within each
176 technical replicate (Port et al., 2016). To derive this index, we first normalized taxon-specific
177 ASV counts into proportions within a technical replicate, and then transformed the proportion
178 values such that the maximum across all samples is scaled to 1 for each taxon. Such indexing
179 improves our ability to track trends in abundance of individual taxa in time and space by
180 correcting for both differences in read depth among samples and differences in amplification

efficiency among sequences; mathematically, it is equivalent to the Wisconsin double-standardization for community ecology as implemented in vegan (Oksanen et al., 2013). Using this index, we generated a single Bray-Curtis dissimilarity matrix for sequenced transect samples from each unique site/month combination and performed ordinations for each using the metaMDS function of the vegan package for R (Oksanen et al., 2013; R Core Team, 2016) using a maximum of one-thousand random starts. We then created a single Bray-Curtis dissimilarity matrix for our entire dataset and apportioned variance by site, month, transect distance, and sample on the communities present using a PERMANOVA test (implemented with the adonis function (Oksanen et al., 2013).

Habitat preference

To examine the abundance of sequences from each phylum in eelgrass habitat relative to bare substrate, we first assigned taxonomy to ASVs and trimmed our dataset to taxa within phyla represented by a total of at least 10,000 reads, a natural break in the histogram of read counts (Figure S2). To visualize the ecological patterns across taxa, we then examined eDNA indices for each phylum at the two transect extremes (within-eelgrass vs. bare), calculating a relative eDNA abundance measure by subtracting the mean eDNA abundance index over bare substrate for each site-month combination from the corresponding mean eDNA abundance index in the eelgrass habitat. Positive values of this measure thus denote higher abundance in eelgrass, while negative values of this index indicate higher abundance over bare substrate. To assess the statistical significance of these phylum-level differences between habitat types, we compared the distributions of mean eDNA abundance indices for individual phyla in samples taken from eelgrass relative to their counterparts taken over bare substrate, using a paired Wilcoxon signed rank test with Bonferroni correction for multiple comparisons.

Dinoflagellate transect patterns

To examine dinoflagellate abundance patterns along the transects from eelgrass to bare habitat, we first assigned taxonomy to family (or genus, when possible) for all ASVs with at least 1000 reads from the phylum Dinoflagellata. We chose to consider only dinoflagellates with high read counts in our dataset not necessarily because they are the most prevalent in the environment (raw read counts are biased by differences in amplification efficiency), but because substantial numbers of reads allow us to draw more robust conclusions about the distribution of taxa at the fine geographic scale of our transects. We again calculated an eDNA abundance index for all technical replicates of each available biological sample on the alongshore transects between the two habitat extremes. Because plotting data for individual taxa across transects for each site-month revealed extremely episodic abundance of dinoflagellate sequences (Figure S3), we used the k-means function of the R stats package (Team & Worldwide, 2015) to separate high- and low-abundance transects across all dinoflagellate taxa with unsupervised machine learning (Figure S5). Specifically, we took the grand mean of taxon-specific eDNA indices for each technical replicate along transects at a given time and place, and subjected these values to clustering with two groups ($k = 2$).

Eight transects identified by unsupervised clustering indicate high-abundance events within at least one taxon. For these focal transects, we first compared the eelgrass and bare habitat using a paired Wilcoxon signed-rank test of mean eDNA abundance index for each dinoflagellate taxon (here, having identified sequences to the level of family or genus, rather than grouping dinoflagellates together, as we have done above) . Next, to determine whether dinoflagellate abundance measures at intermediate alongshore transect samples (1, 3, 6, 10, and 15 meters) were more closely associated with eelgrass or bare habitat, we additionally performed Gaussian

mixture modelling with two groups (Scrucca et al., 2016). We then used a Wilcoxon rank sum test to assess the significance of differences in the dinoflagellate eDNA abundance index distribution in the two groups produced by model-based clustering. To ensure that these groups did not result simply from spatial autocorrelation, we calculated Bray-Curtis dissimilarity based on eDNA abundance indices of all ASVs from adjacent points on each full transect. We tested the null hypothesis that spatial distance does not significantly influence Bray-Curtis dissimilarity using a Kruskall-Wallace test.

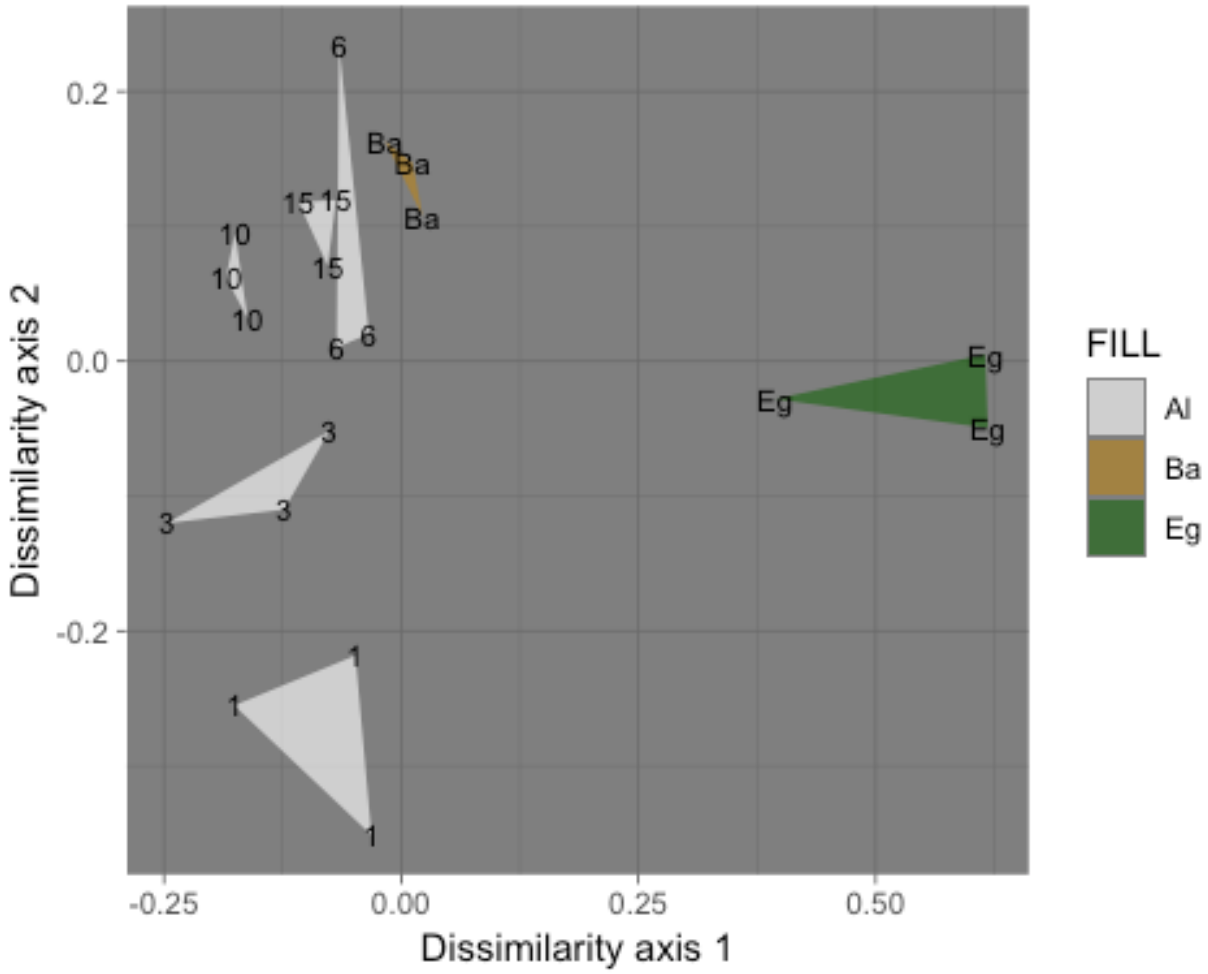
Results

Community Composition

We assigned over 3,000 unique ASVs to 12 phyla comprising a diverse set of single- and multicellular taxa including Arthropoda (arthropods), Annelida (annelid worms), Bacillariophyta (diatoms), Bacteriodetes (division of gram-negative, rod-shaped bacteria), Chlorophyta (green algae), Chordata (chordates), Cnidaria (cnidarians), Dinoflagellata (dinoflagellates), Echinodermata (echinoderms), Mollusca (molluscs), Ocrophyta (brown algae), and Rhodophyta (red algae). This represents a broad – although by no means comprehensive – survey of eukaryotic communities in and around our sampled eelgrass beds.

nMDS ordination revealed consistent differentiation between eDNA communities across transects within a sampling site and date; technical replicates consistently clustered together. An example plot of samples gathered along the transect from eelgrass to bare substrate at Willapa Bay in July (Figure 2; all site/date plots shown in Figure S7) shows that the eelgrass community is quite dissimilar from other transect points along both axes. Moving away from eelgrass, all

three technical replicates of each sample bottle are fully distinguishable from those of other sample bottles (non-overlapping in ordination). For the instances in which complete transects were sampled at a given time and place (10) and all three technical replicates of a sample were available for analysis (60), 47 samples (78%) were similarly non-overlapping in ordination with all remaining transect points, demonstrating that despite proximity at the scale of meters, bottles of water contained eDNA evidence of distinct biological communities the majority of the time. Put differently, within-sample variance (reflecting laboratory-driven processes) was smaller than between-sample variance (reflecting biological as well as laboratory processes), hence providing resolution of communities at the scale of meters.



257

258 *Figure 2: Example ordination plot of samples along a single transect from bare to eelgrass*
259 *positions at Willapa Bay in July, 2017. Technical replicates of each biological sample are*
260 *grouped as triangles. White alongshore transect samples are labeled with distance from eelgrass*
261 *in meters; the single within-bed sample is green (labeled Eg) and the bare sample is brown (Ba).*

262 PERMANOVA apportioned the variance in Bray-Curtis distance among samples as follows: site
263 ($R^2 = 0.18593$, $p = 0.001$), month ($R^2 = 0.07909$, $p = 0.001$), and transect distance ($R^2 =$
264 0.02625, $p = 0.001$) each explain a significant portion of the variance in the dataset. Thus,
265 despite strong effects of geographic location and season, we do see a highly significant effect of
266 proximity to eelgrass on the complement of organisms present. Moreover, these results confirm

267 that we can consistently distinguish nearshore eDNA communities – as sampled by our primers –
268 at spatial scales of meters.

269 Habitat Preference

270 To determine the habitat preference of major taxa in our dataset at a course spatial scale, we
271 classified ASVs to the level of phylum and plotted an index of their relative sequence abundance
272 in eelgrass versus bare positions (Figure 3). Positive indices denote greater abundance in
273 eelgrass, and negative indices in bare substrate. Across all sites and months, only dinoflagellates
274 show a consistent and strong bias towards one habitat or another; they are nearly universally
275 more abundant in bare habitat. Indeed, the negative association of dinoflagellates with eelgrass
276 beds is the only significant result of tests of phylum abundance in the two habitat extremes after
277 Bon Ferroni correction for multiple comparisons ($a = 0.0042$, $p = 0.004$; paired Wilcoxon signed
278 rank test). Other single-celled microalgae such as diatoms (Bacillariophyta) and green algae
279 (Chlorophyta) do not show these same patterns of distribution with respect to eelgrass.

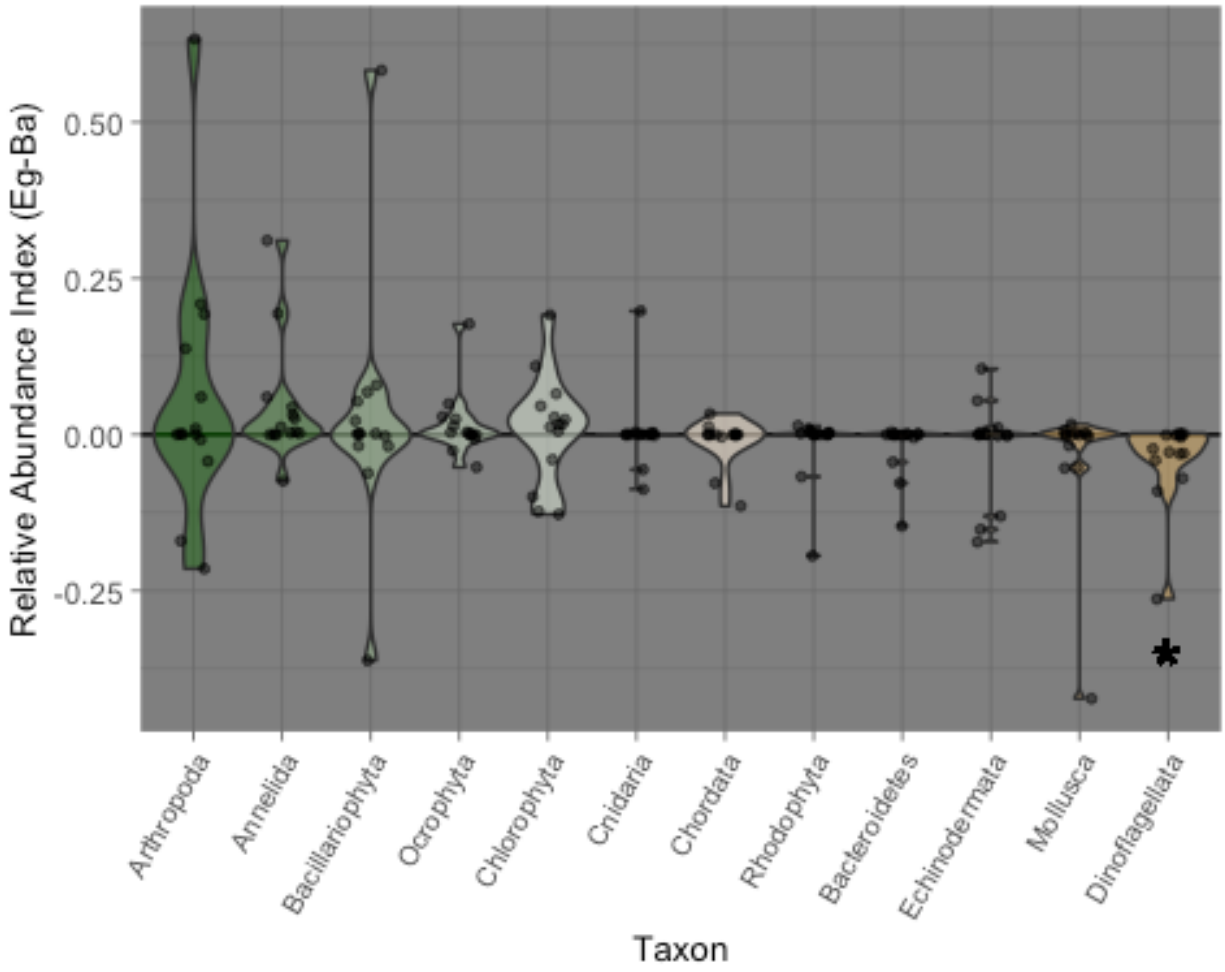


Figure 3: Habitat preferences of sequences within each phylum. Phyla are ordered and colored by mean relative abundance index (eDNA abundance index in eelgrass - eDNA abundance index over bare substrate). Greener samples on the left exhibit greater relative abundance in eelgrass, and browner samples on the right exhibit greater relative abundance on bare substrate.The central zero-line indicates no bias in abundance between habitat types.

Dinoflagellate Distributions

To assess the patterns of dinoflagellate abundance that contribute to an overall preference for habitat bare of eelgrass, we first honed our focus to the eleven dinoflagellate ASVs - roughly,

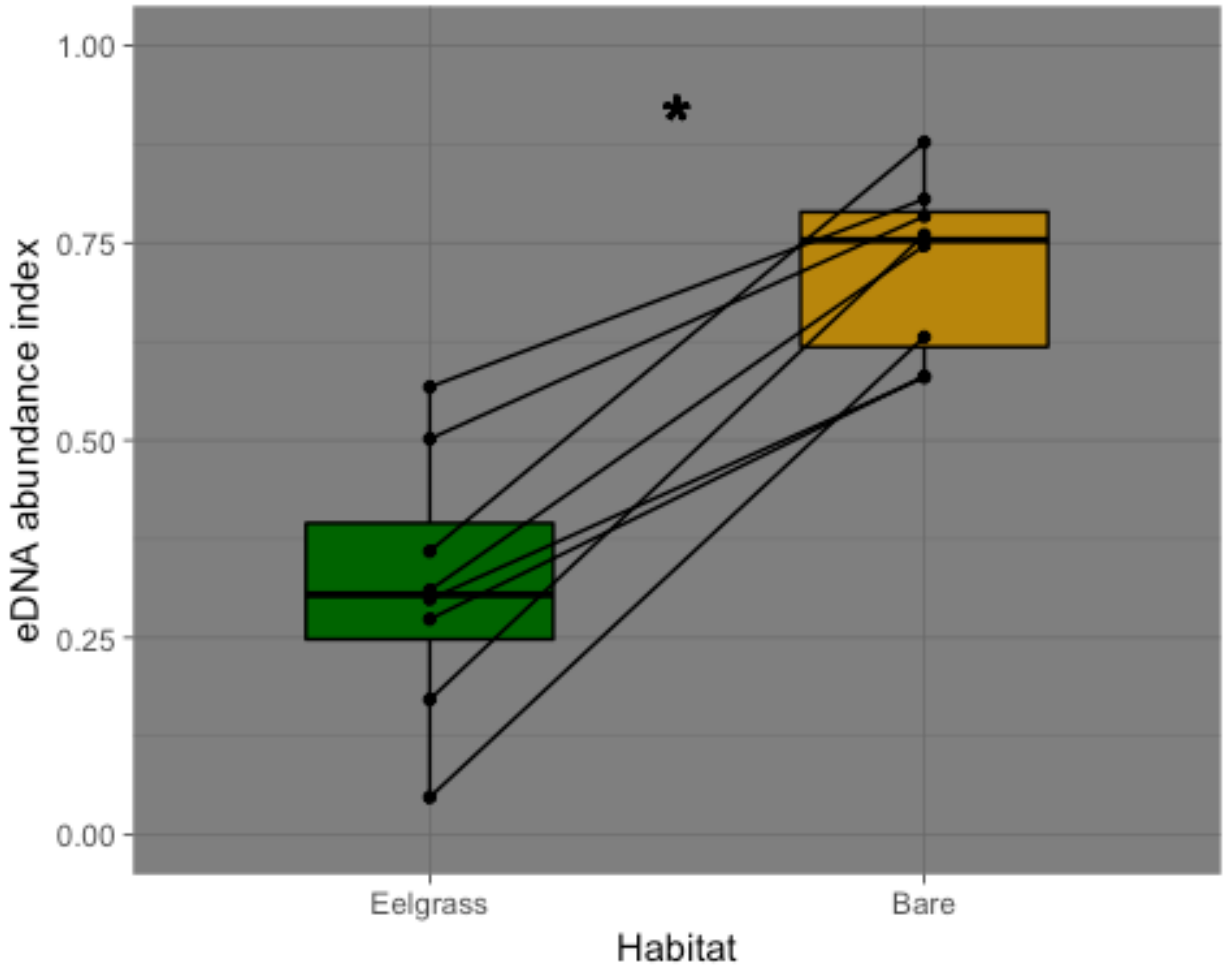
289 species - with more than 1000 reads in our dataset (Table 2). eDNA indices suggest that
290 dinoflagellate distributions are highly local and episodic at the scale of our sampling (Figure S3);
291 each dinoflagellate taxon appears at appreciable levels at only a single site and in no more than
292 two consecutive sampling periods in our dataset. It is when dinoflagellates are plentiful relative
293 to background levels that we have the power to identify trends in the abundance of individual
294 taxa with respect to eelgrass habitat. To restrict our analysis to such periods, we used
295 unsupervised machine learning (k-means clustering) to define a set of high- and low-abundance
296 transects for each dinoflagellate sequence across all sites and months (Figure S5; between group
297 sum of squares / total sum of squares = 81.1 %); eight transects from seven dinoflagellate taxa
298 appeared in the high-abundance group.

299 *Table 1: Taxon (given as Family (Genus)) and total sequence read count for each dinoflagellate*
300 *ASV with >1000 total sequence reads.*

Taxon	Count
Gonyaulacaceae (Alexandrium)	47010
Heterocapsaceae (Heterocapsa 1)	29912
Gonyaulacaceae (Protoceratium)	11948
Gymnodiniaceae (Nusuttodinium)	8096
Heterocapsaceae (Heterocapsa 2)	7382
Gonyaulacaceae (unknown)	3897
Kareniaceae (Karlodinium)	3515
Gymnodiniaceae (Gymnodinium)	3249

Kareniaceae (unknown)	2521
Syndiniaceae (Hematodinium)	1401
Peridiniales (Protoperidinium)	1278

In this subset of high-abundance transects, we observe that the negative interaction of eelgrass and dinoflagellates is taxonomically universal; all sequences (from families Gonyalacaceae, Heterocapsaceae, and Kareniaceae, each of which include known or suspected HAB species (IOC Harmful Algal Bloom Programme and the World Register of Marine Species)) are heavily biased towards bare substrate, relative to eelgrass (Figure 4). A comparison of the mean eDNA abundance index across technical replicates for each high-abundance transect taxon demonstrates that this preference for bare substrate over eelgrass is significant (Wilcoxon signed-rank test, $p < 0.008$).



309

310 *Figure 4: Habitat preferences of dinoflagellate sequences at site-months in which each taxon*
311 *occurs at high abundance.*

312 After demonstrating a preference of all dinoflagellate taxa towards the bare habitat extreme
313 (when highly abundant), we then characterized the possible influence of eelgrass on the
314 immediately surrounding environment, as a function of distance from the edge of the beds, using
315 data from entire transects (Figure 5). Examining all points alongshore, we found that
316 dinoflagellate eDNA abundance indices at the 1, 3, 6, 10, and 15m positions grouped with those
317 at the eelgrass position in model-based clustering (probability of assignment to the group with
318 eelgrass samples was in each case at least 10^34 more likely than probability of assignment to

the group with bare samples). Additionally, the eDNA abundance index of all high-abundance dinoflagellate taxa at these six transect points together differed significantly from bare substrate (Wilcoxon signed rank test, $p < 0.02$). These patterns are not simply due to spatial autocorrelation, as overall Bray-Curtis dissimilarity (from all ASVs) shows no pattern associated with geographic distance across full transects (Figure S6; Kruskall-Wallis rank sum test, $p > 0.9$).

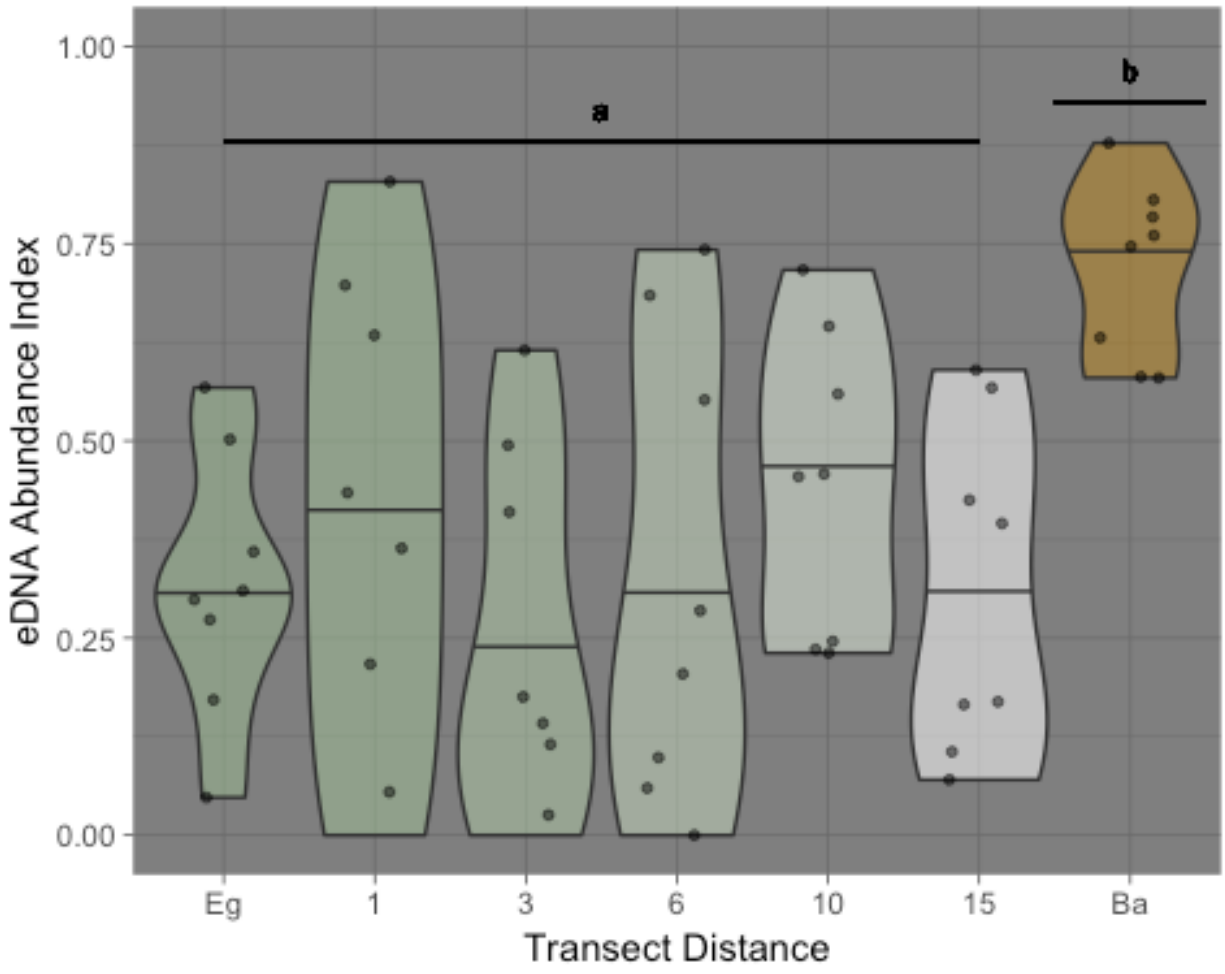


Figure 5: Dinoflagellate eDNA abundance measures plotted for all sites and months combined at each point along the transect from eelgrass to bare substrate, with median shown.

Discussion

In a broad-spectrum eDNA survey of the organisms living in and near to eelgrass, we track the relative abundance of a diverse group of taxa representing twelve phyla. We demonstrate the ability of eDNA to distinguish communities represented in samples taken only meters apart, and to reveal a significant axis of variance based on proximity to habitat type, despite strong influences of geography and season across sampling events. One major and significant pattern emerges in our analysis: highly-abundant dinoflagellate taxa are more common over bare substrate than within eelgrass beds, and the putative effect of eelgrass extends at least 15m beyond the edge of the beds themselves.

In line with our community-level observation, a specific allelopathy against microalgal species by *Z. marina* was first described over 30 years ago (Harrison & Durance, 1985). More recent evidence suggests that this negative interaction applies to multiple HAB taxa (including *Alexandrium*, also observed in this study), and is mediated in our sampling locations by a variety of strains of eelgrass-associated algicidal and growth-inhibiting bacteria, particularly from Erythrobacter, Teredinibacter, Gaetbulibacter, and Arthrobacter genera (Inaba et al., 2017) (though the eDNA primers employed here amplify eukaryotes almost exclusively and therefore do not allow us to test this mechanism directly). However, in our dataset the repressive effect of eelgrass notably does not extend at the phylum level to other phytoplankton such as diatoms (*Bacillariophyta*) and green algae (*Chlorophyta*), despite reports that *Z. marina* habitat can deter members of these taxa as well (reviewed in Gross, 2003).

Dinoflagellates responsive to eelgrass habitat when at high abundance in our dataset include species from the genera *Heterocapsa*, *Alexandrium*, *Karlodinium*, *Protoceratium*, and from

families *Gonyaulacaceae* and *Kareniaceae*, each of which have at least one member included in local microscopy-based monitoring programs (Amelia Kolb & Swanson, 2016; Vera Trainer, 2016); our eDNA methodology thus agrees broadly with previous visual identification of microalgae. Of particular interest are dinoflagellate taxa that include HAB-forming members: the resident species of *Alexandrium* (*A. catanella*) causes paralytic shellfish poisoning via production of saxitoxin (STX; Wiese et al., 2010), and species from both the genus *Protoceratium* (e.g. *P. reticulatum*) and the family *Gonyaulacaceae* (e.g. *Gonyaulax spinifera*) produce yessotoxins (YTXs), whose effects on human consumers of contaminated shellfish are complex and unclear (reviewed in Tubaro et al., 2010). Toxins from these three taxa impact the aquaculture and harvest industries directly; detection of STX at concentrations greater than 80 µg STXequiv/100 g is routinely responsible for regional harvest closures (Moore et al., 2009), and shellfish containing more than 0.1 µg YTX equiv/100g may not be sold to markets within the European Union, although this toxin is not currently regulated within the US (Trainer et al., 2013). In summary, the dinoflagellate taxa deterred by eelgrass habitat in this study have high relevance for local shellfish management decisions, particularly as HABs (including *Alexandrium*) are intensifying with recent ocean warming in the North Pacific (Gobler et al., 2017).

In order to understand the relationship of *Z. marina* to ecosystem and human health, as well as to shellfish farming and harvest, it is critical to consider our addition of an 'action-at-a-distance' element to the existing eelgrass-dinoflagellate interaction model. Given the protected status of *Z. marina* habitat on the Pacific Coast of the United States, the goals of the shellfish industry and eelgrass conservation are often perceived as being in conflict (Forrest et al., 2009) and policies prohibit shellfish farming and harvesting within or near beds. For example, in Washington State,

373 required buffer zones between shellfish aquaculture and eelgrass range from 3 to 8m, depending
374 on the agency involved (National Marine Fisheries Service West Coast Region, 2017). However,
375 our work demonstrates that *Z. marina* habitat may have a protective effect against harmful
376 dinoflagellates within these buffer zones, reducing the potential for shellfish to accumulate HAB
377 toxins from the surrounding waters. Likewise, filter feeders can mitigate microbial disease in
378 adjacent environments, and *M. gigas*, in particular, has recently been shown to lessen the effects
379 of Eelgrass Wasting Disease (EWD) on *Z. marina* (Groner et al., 2018). As others have begun to
380 suggest, then, eelgrass and oysters may be critical allies to one another in changing marine
381 ecosystems worldwide. Future work will examine their multi-faceted symbiosis, particularly in
382 characterizing the taxonomic breadth of potential seagrass and shellfish partnerships, as well as
383 in defining the molecular mechanisms underlying the roles of both beneficial and detrimental
384 microbial intermediates.

385 ## Supplemental Materials

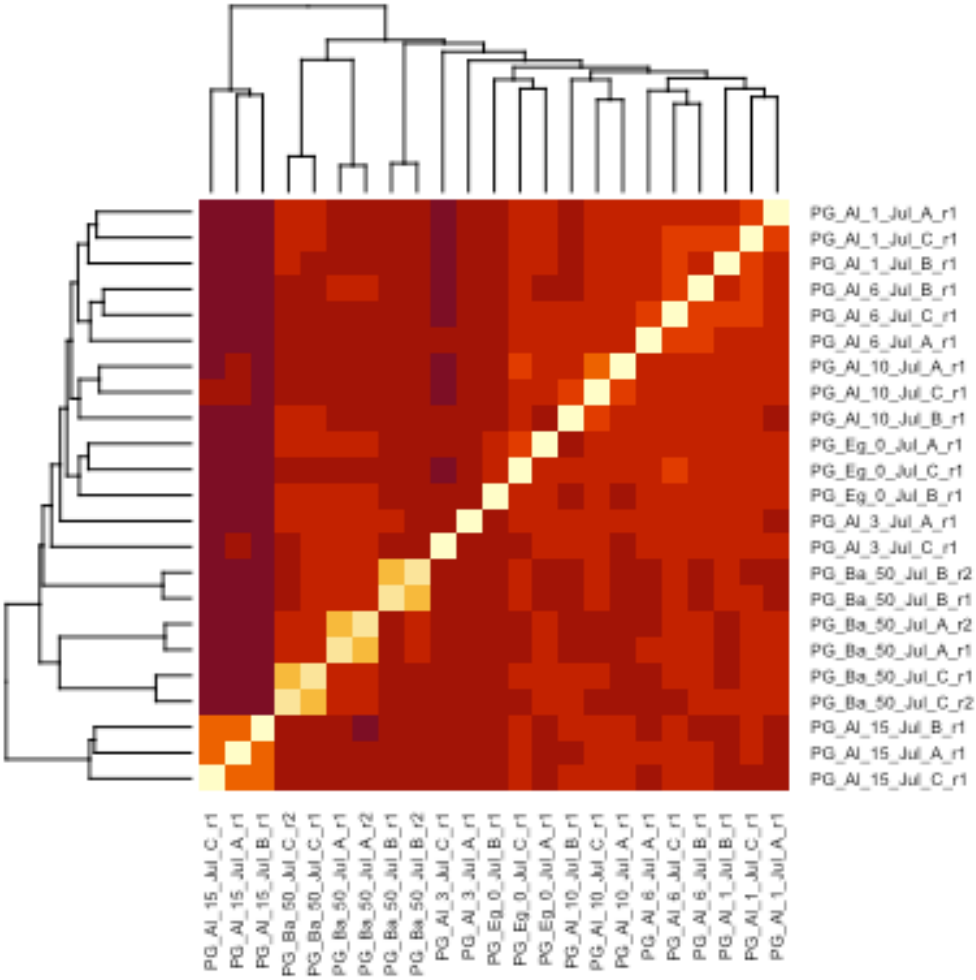
386 *Table S1: Sample information. For each site, Case Inlet (CI), Port Gamble (PG), Nisqually*
387 *Reach (NR), Skokomish (SK), and Willapa Bay (WB), approximate transect positions are*
388 *recorded, as well as latitude, longitude, and the approximate geographic distance of each*
389 *sample from the eelgrass bed edge, calculated from coordinates. Negative distances indicate*
390 *samples within the eelgrass bed itself. Columns named May, July, and August list the number of*
391 *technical replicates passing quality control measures of three sequenced from each bottle of*
392 *water. NA indicates samples that were not gathered, and asterisks indicate samples for which*

393 *three technical replicates were sequenced on two separate MiSeq runs to characterize the*

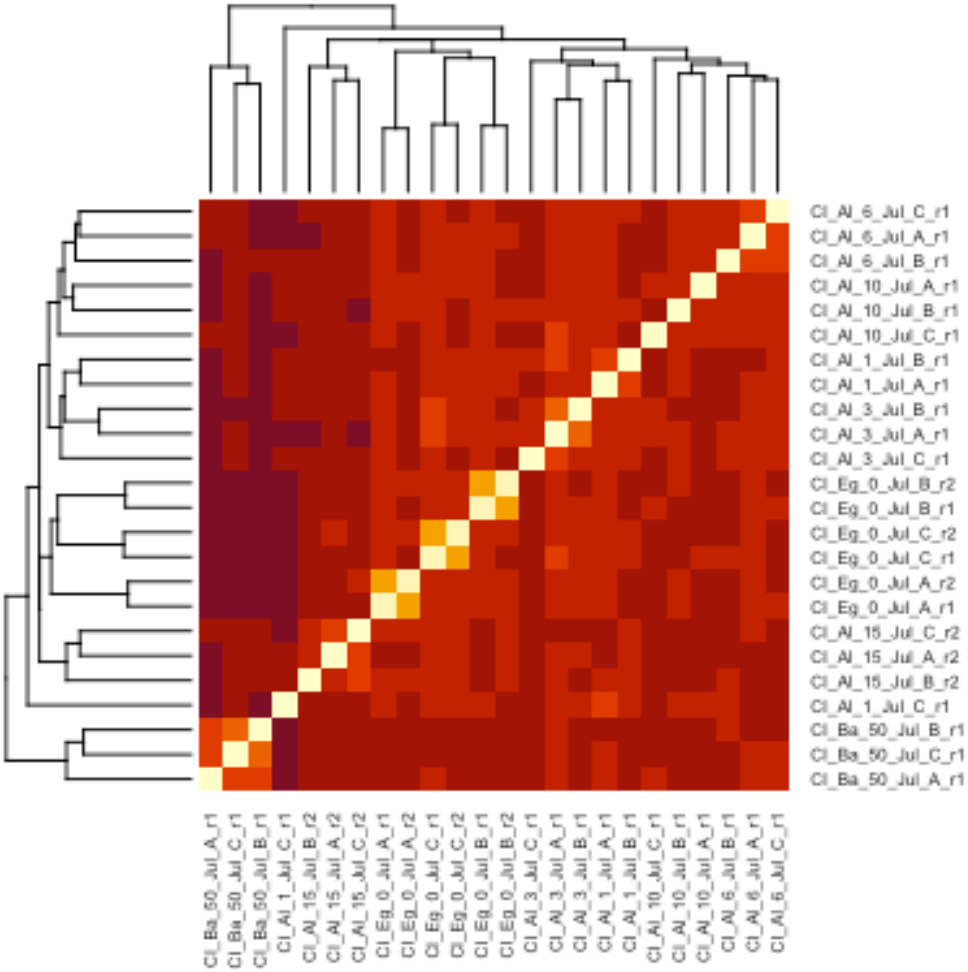
394 *importance of sequencing run in explaining variation among samples.*

Site	Position	long	lat	Distance	May	July	August
CI	Eelgrass	-122.79645	47.358439	-47	3	6*	3
CI	Along 1	-122.79584	47.358455	1	3	3	3
CI	Along 3	-122.796038	47.358565	3	3	3	3
CI	Along 6	-122.795971	47.358551	6	3	3	2
CI	Along 10	-122.795894	47.358481	10	3	3	2
CI	Along 15	-122.795817	47.358436	15	1	3	2
CI	Bare	-122.79576	47.357937	57	3	3	3
NR	Eelgrass	-122.726752	47.101926	NA	3	3	2
NR	Bare	-122.726386	47.101713	NA	3	3	3
PG	Eelgrass	-122.58292	47.847983	-80	3	3	3
PG	Along 1	-122.583221	47.84866	1	3	3	2
PG	Along 3	-122.583157	47.848705	3	3	2	2
PG	Along 6	-122.583222	47.848725	6	3	3	0
PG	Along 10	-122.583278	47.848756	10	3	3	3
PG	Along 15	-122.583258	47.848781	15	3	3	3
PG	Bare	-122.58383	47.842676	666	3	6*	3
SK	Eelgrass	-123.156623	47.354332	-52	3	3	3

SK	Along 1	-123.157147	47.354626	1	NA	3	3
SK	Along 3	-123.157132	47.354585	3	NA	3	2
SK	Along 6	-123.157116	47.354634	6	NA	3	3
SK	Along 10	-123.157162	47.354644	10	NA	3	3
SK	Along 15	-123.157185	47.354733	15	NA	3	2
SK	Bare	-123.157314	47.35502	45	3	3	3
WB	Eelgrass	-124.02619	46.495137	-90	3	3	NA
WB	Along 1	-124.02622	46.494334	1	3	3	2
WB	Along 3	-124.02627	46.494347	3	3	3	3
WB	Along 6	-124.02626	46.494425	6	3	3	3
WB	Along 10	-124.02624	46.494437	10	3	3	3
WB	Along 15	-124.02619	46.494479	15	3	3	3
WB	Bare	-124.026136	46.494479	16	3	3	3

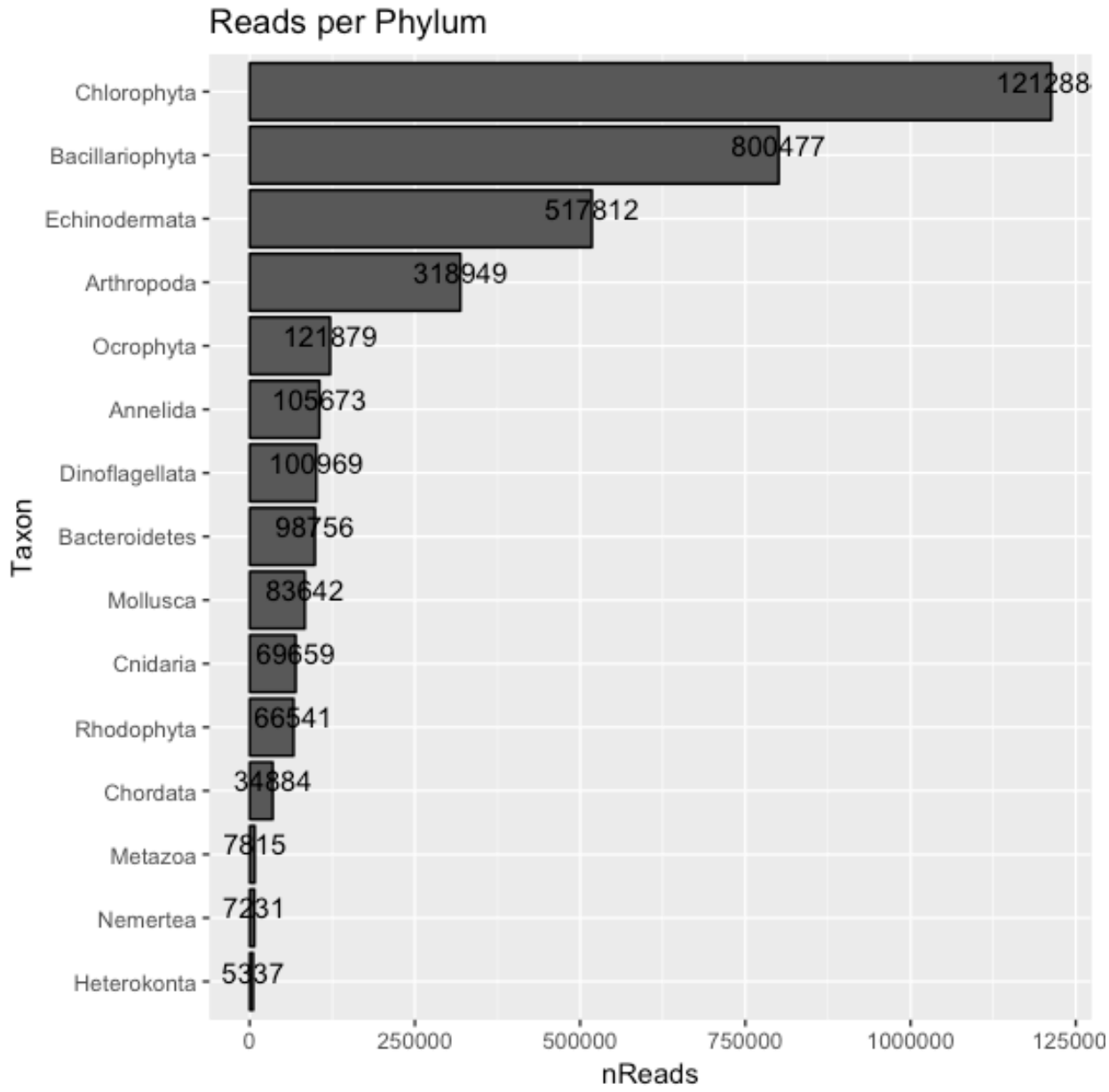


395



396

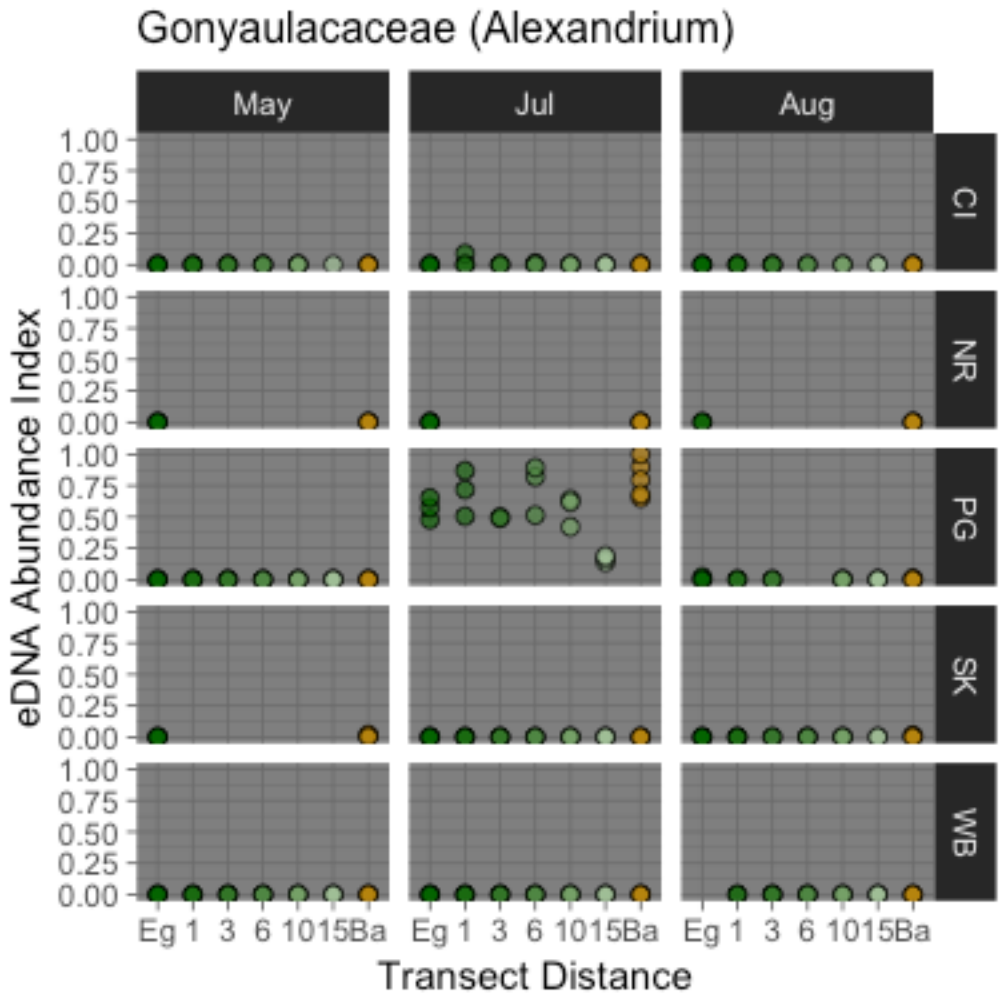
397 *Figure S1: Hierarchical clustering from transect Bray-Curtis distance matrices in which three*
398 *technical replicates were sequenced on two different runs. Names of technical replicates contain*
399 *sample information separated by '_' as follows: Site abbreviation, position abbreviation,*
400 *transect distance, month, replicate, and sequencing run. Note that all replicates from Miseq run*
401 *2 (r2) cluster with the corresponding replicates from Miseq run 1 (r1) for both Port Gamble July*
402 *and Case Inlet July transects.*



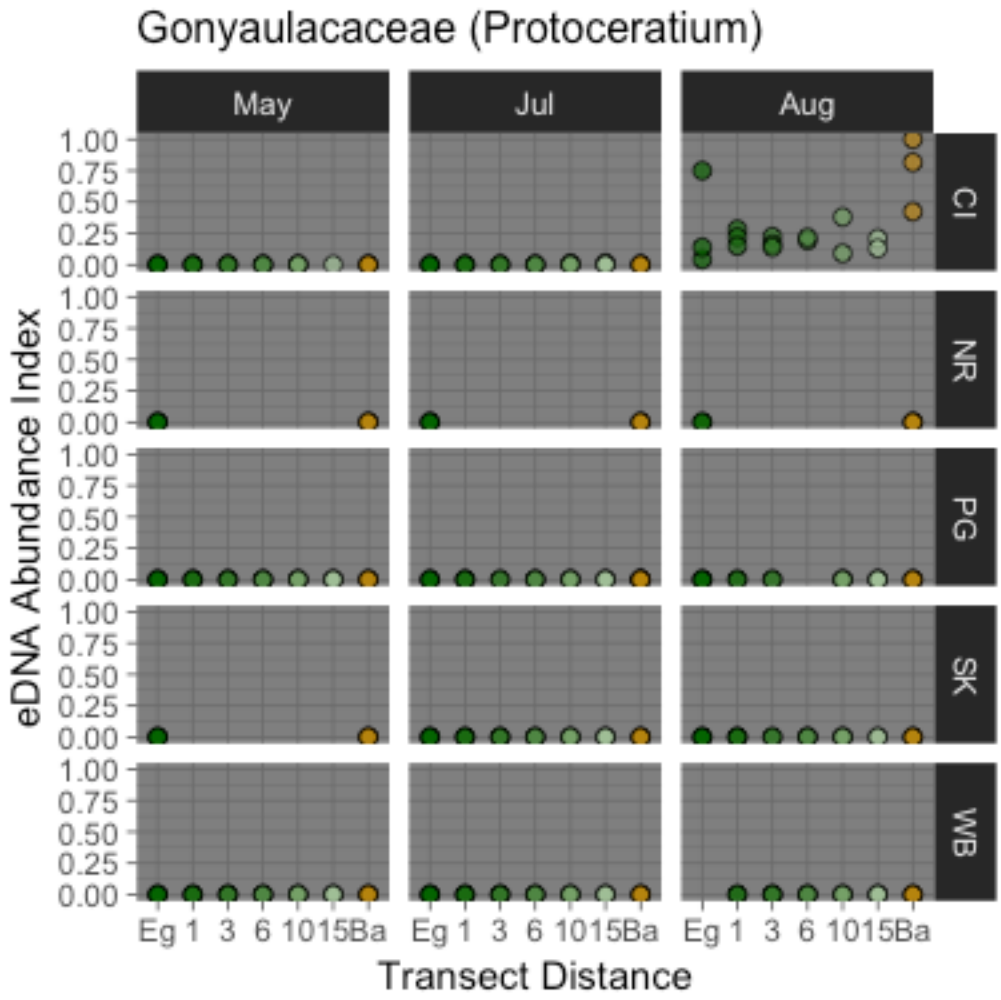
403

404 *Figure S2: Histogram of read counts assigned to each phylum within the complete dataset. Phyla*

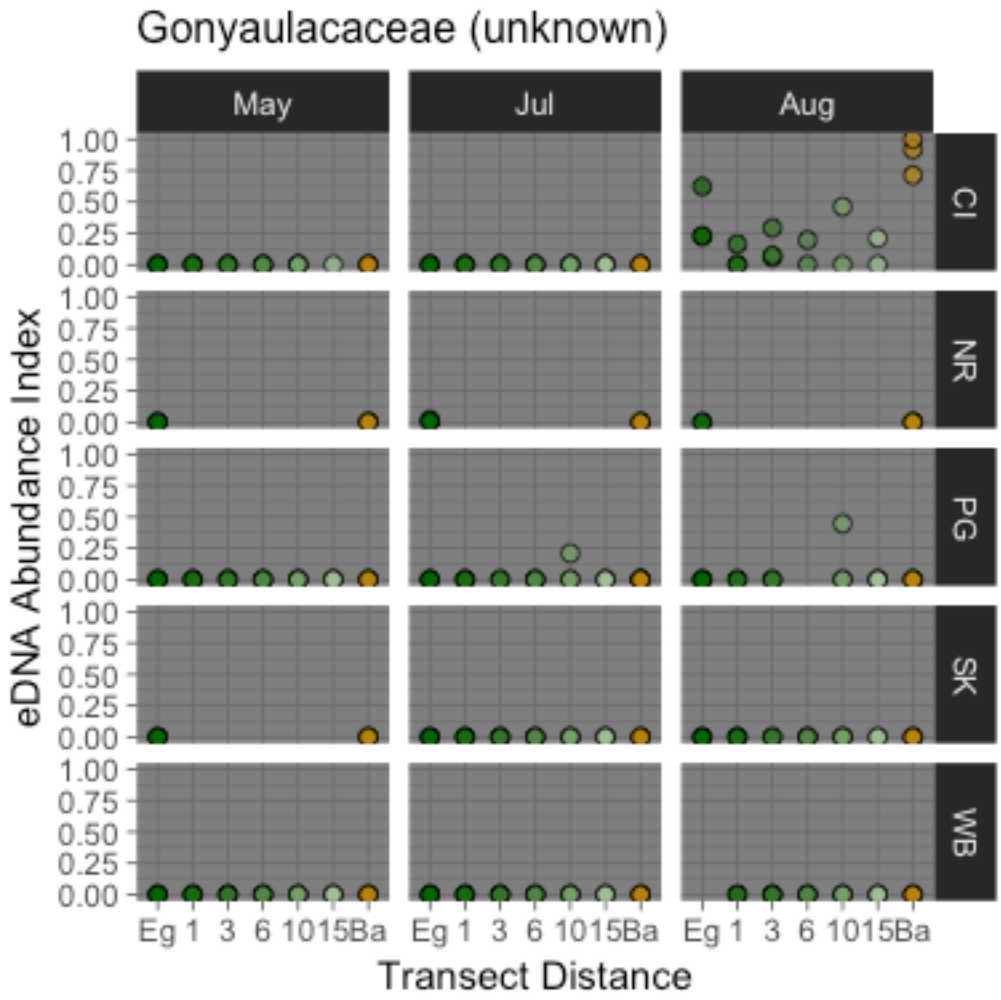
405 *with >10000 reads were chosen for further consideration.*



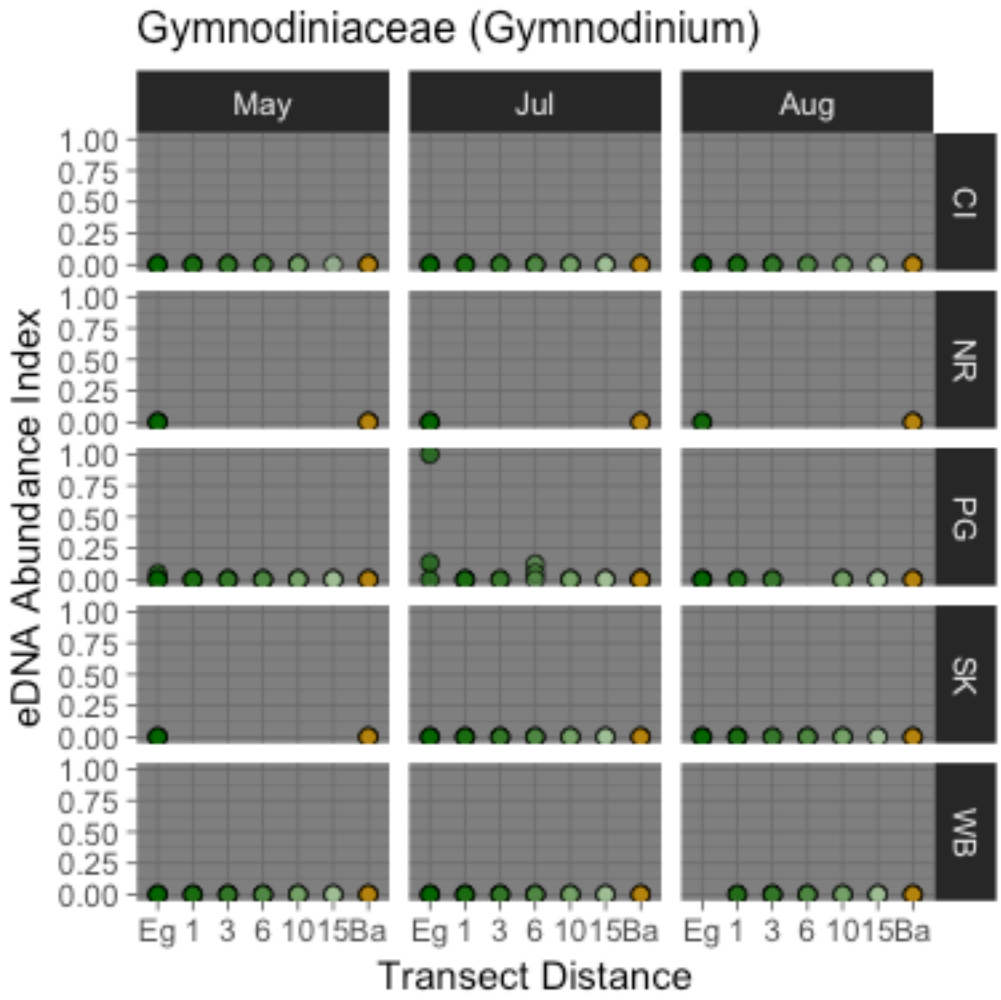
406



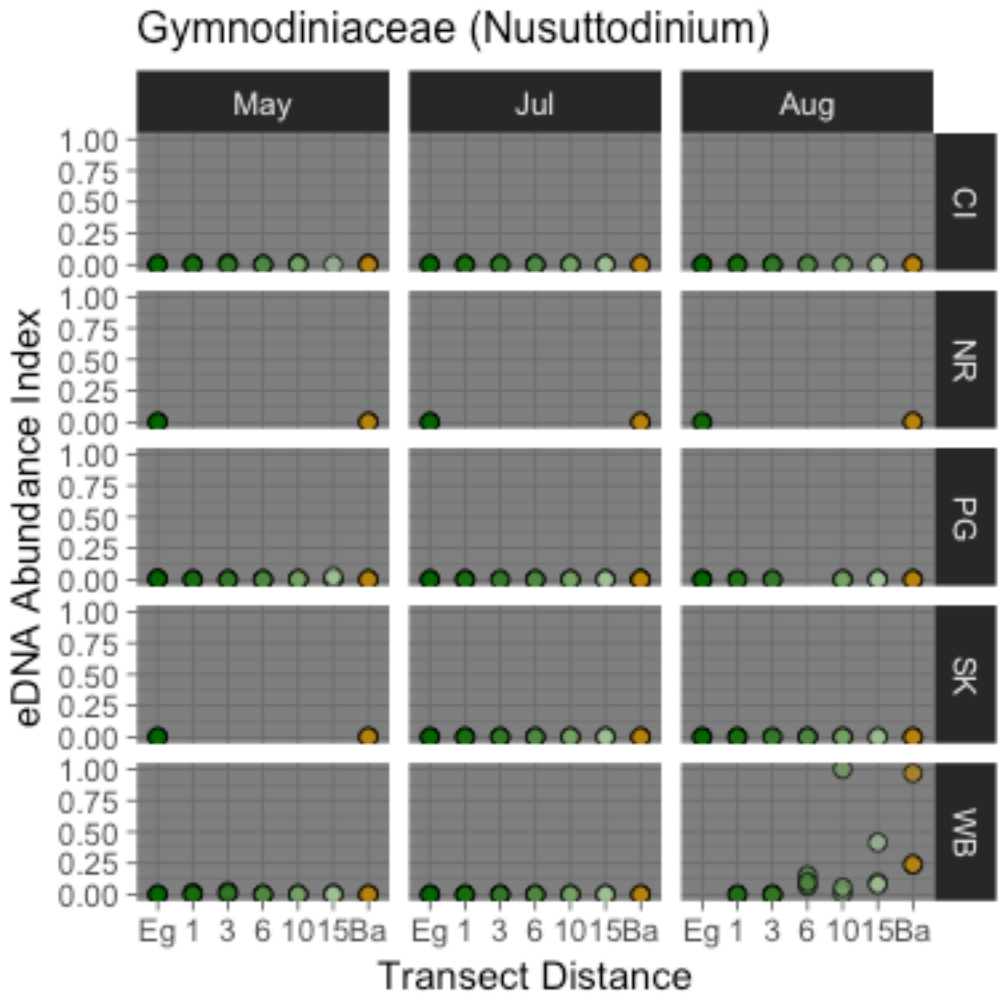
407



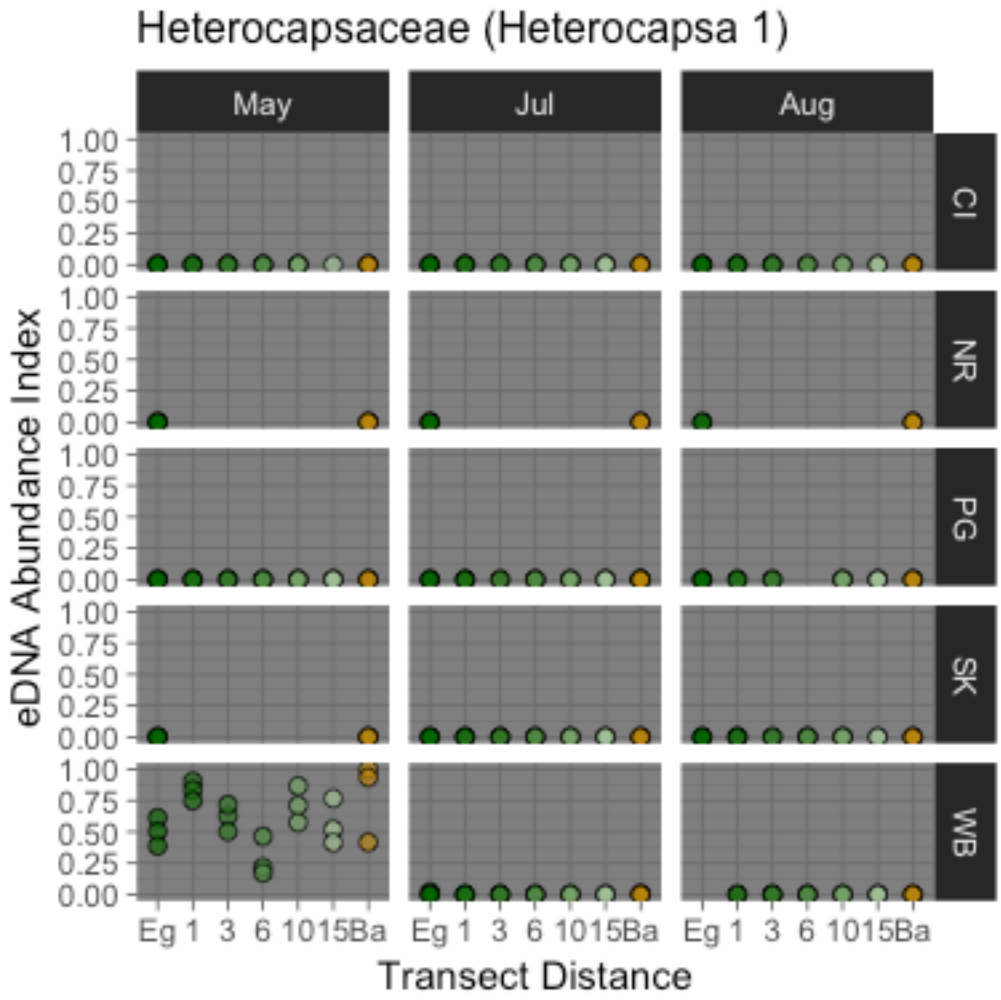
408



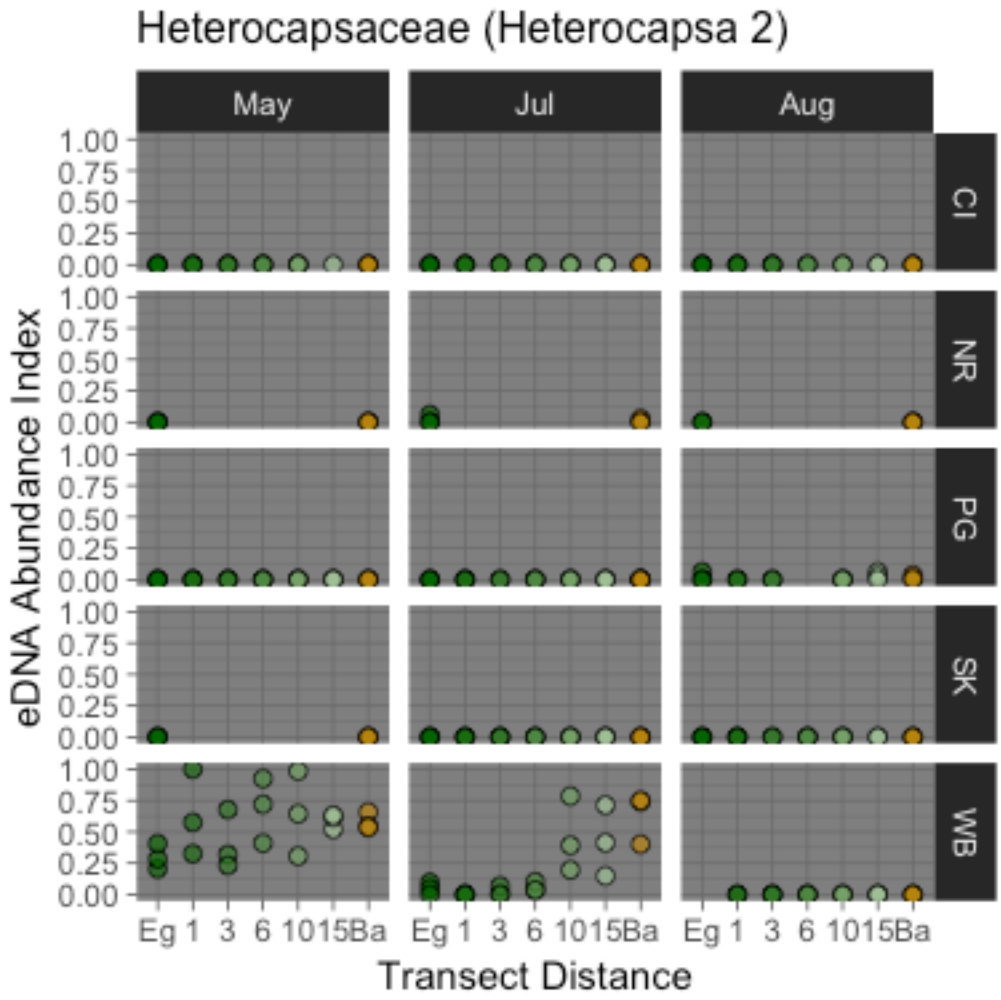
409



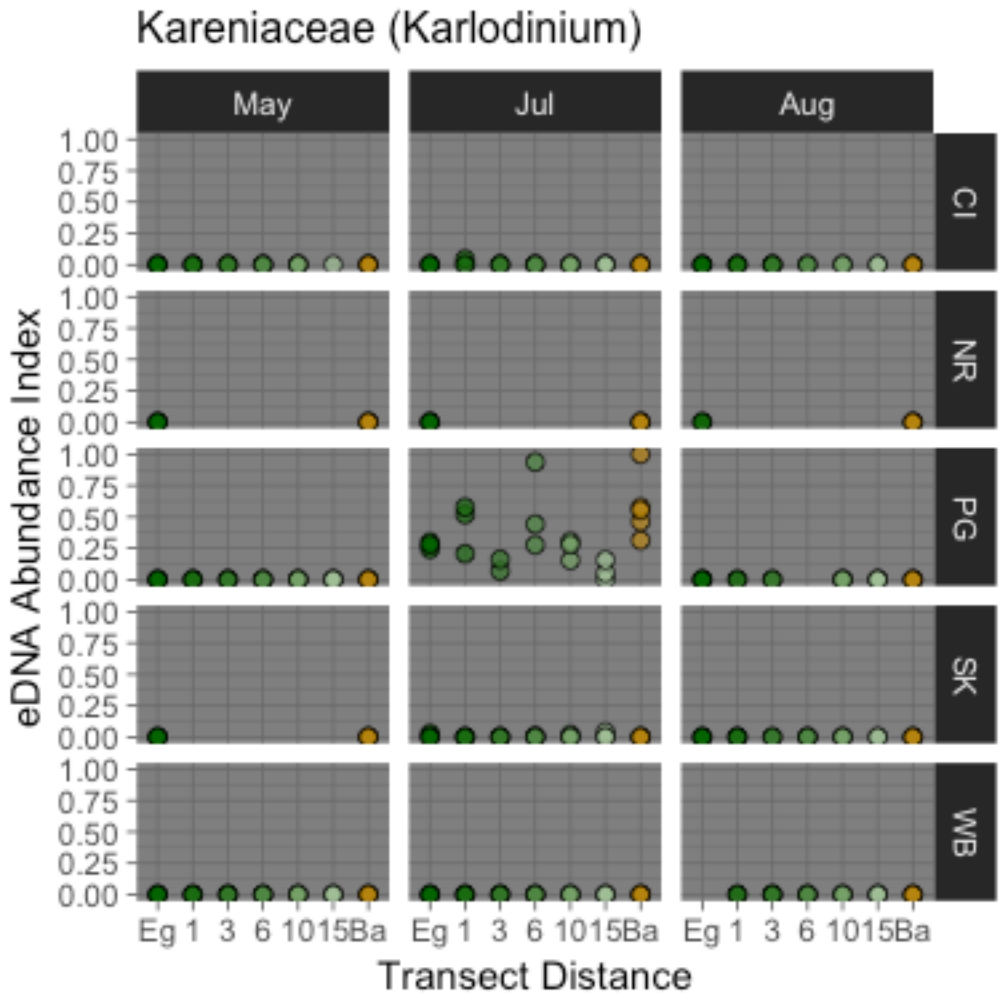
410



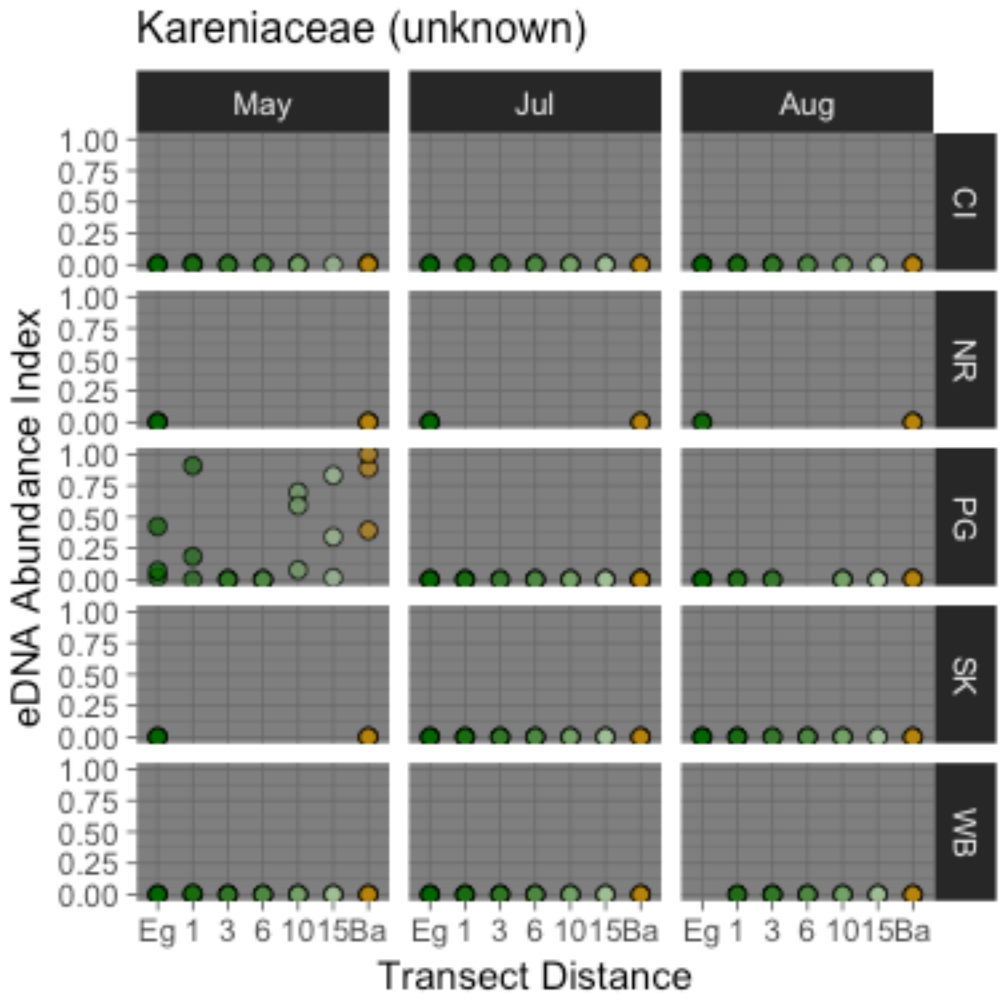
411



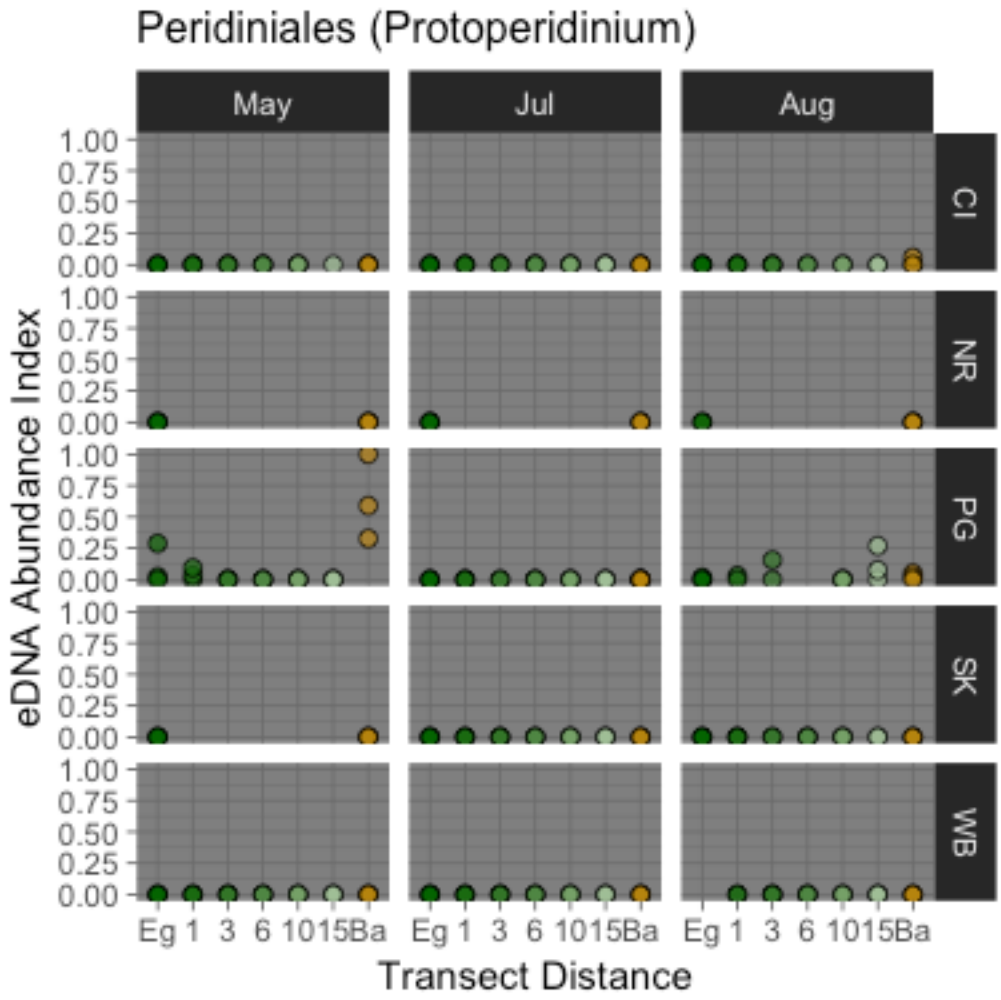
412



413



414



415

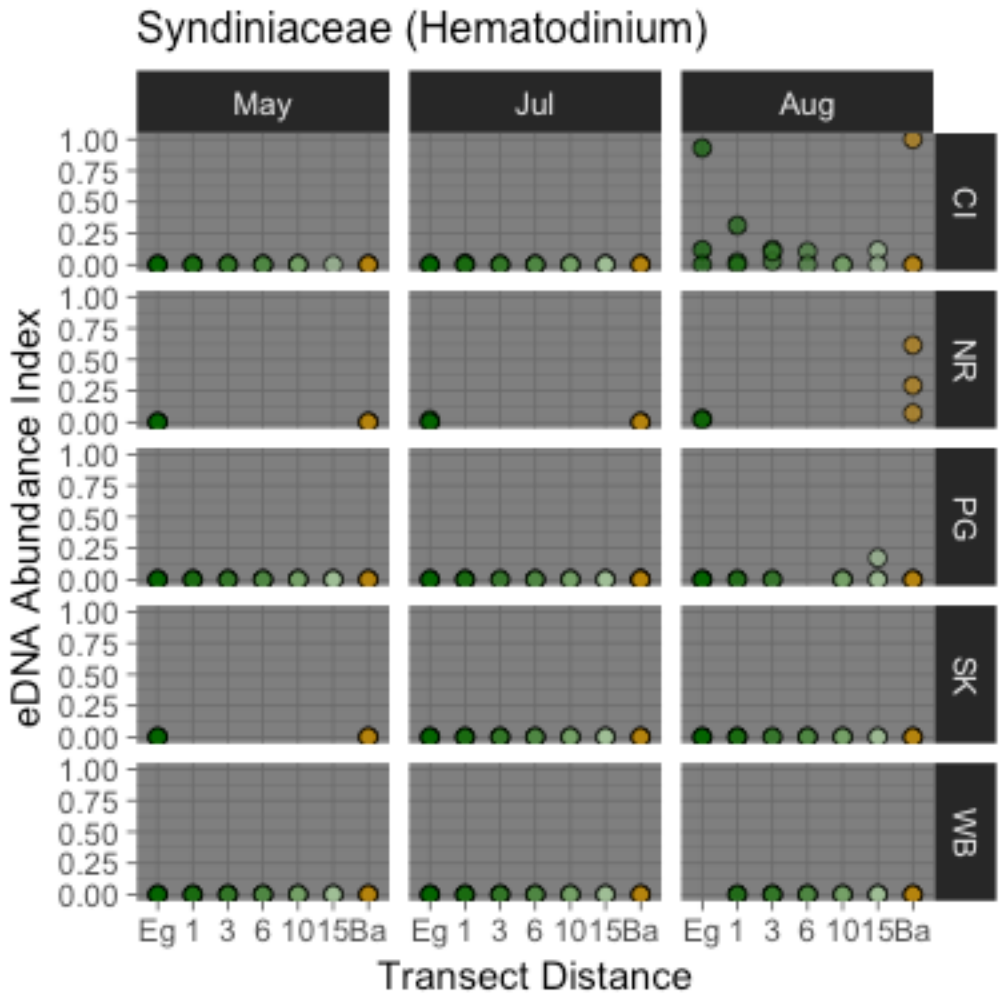
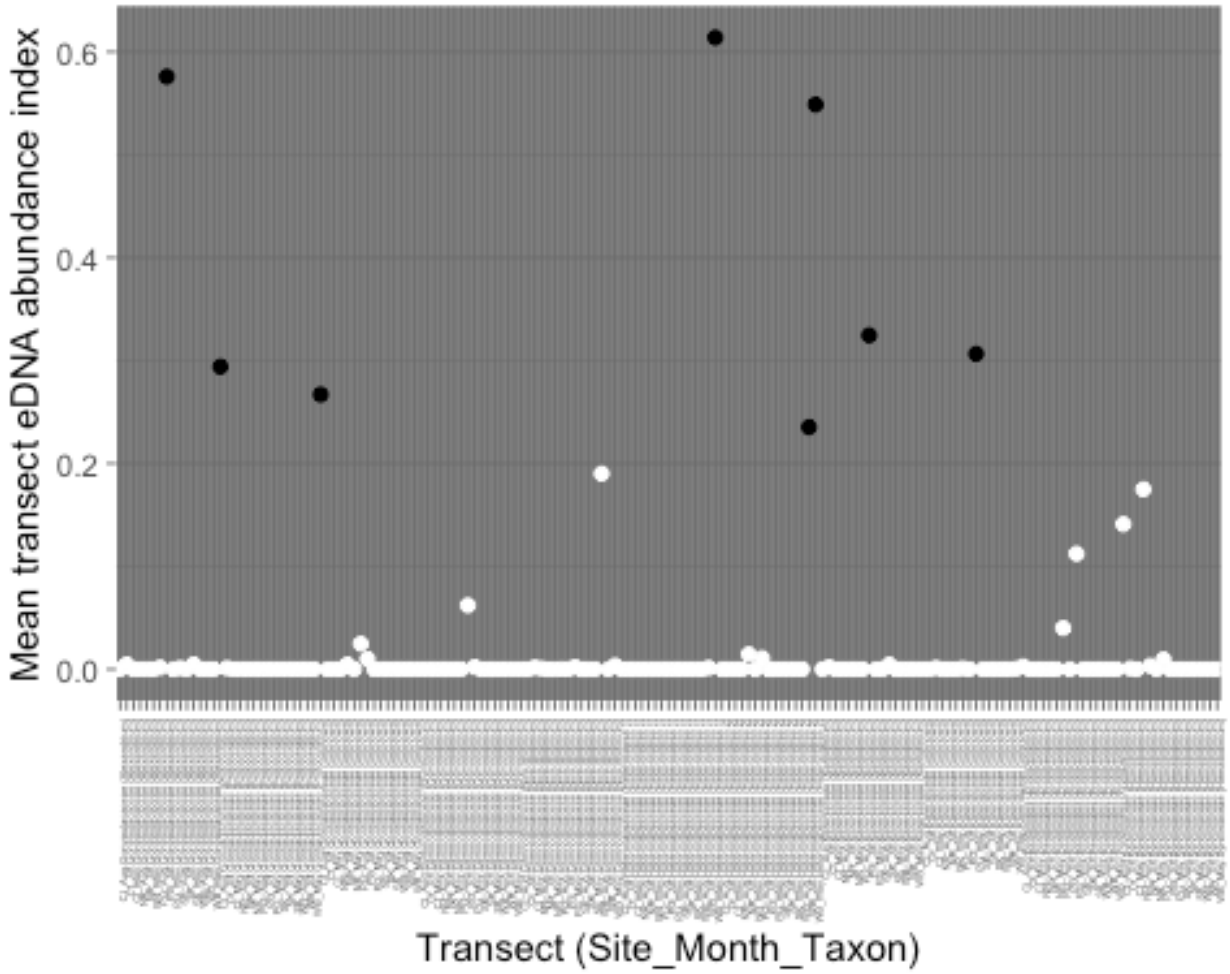


Figure S3: eDNA indices from each technical replicate of biological samples for each dinoflagellate family represented in our dataset by >1000 sequence reads, plotted at all sites for which complete transect data were available. Color indicates proximity to eelgrass habitat (dark green) versus bare substrate (brown).

421

422 *Figure S4: Phylogeny of dinoflagellate sequences from high-abundance transects. Individual*

423 *sequences are named with Family_Genus information (when known).*

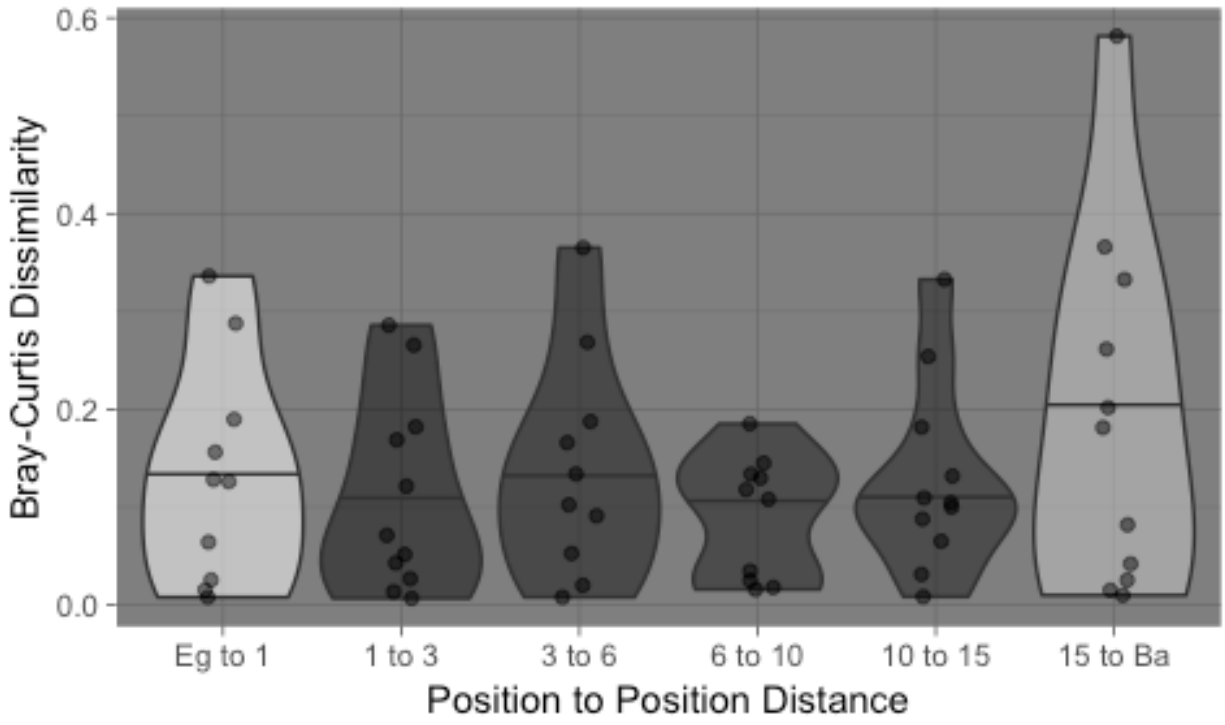


424

425 *Figure S5: High- and low-abundance transects generated with kmeans unsupervised machine*

426 *learning. Black points indicate high-abundance transects; white points indicate low-abundance*

427 *transects*

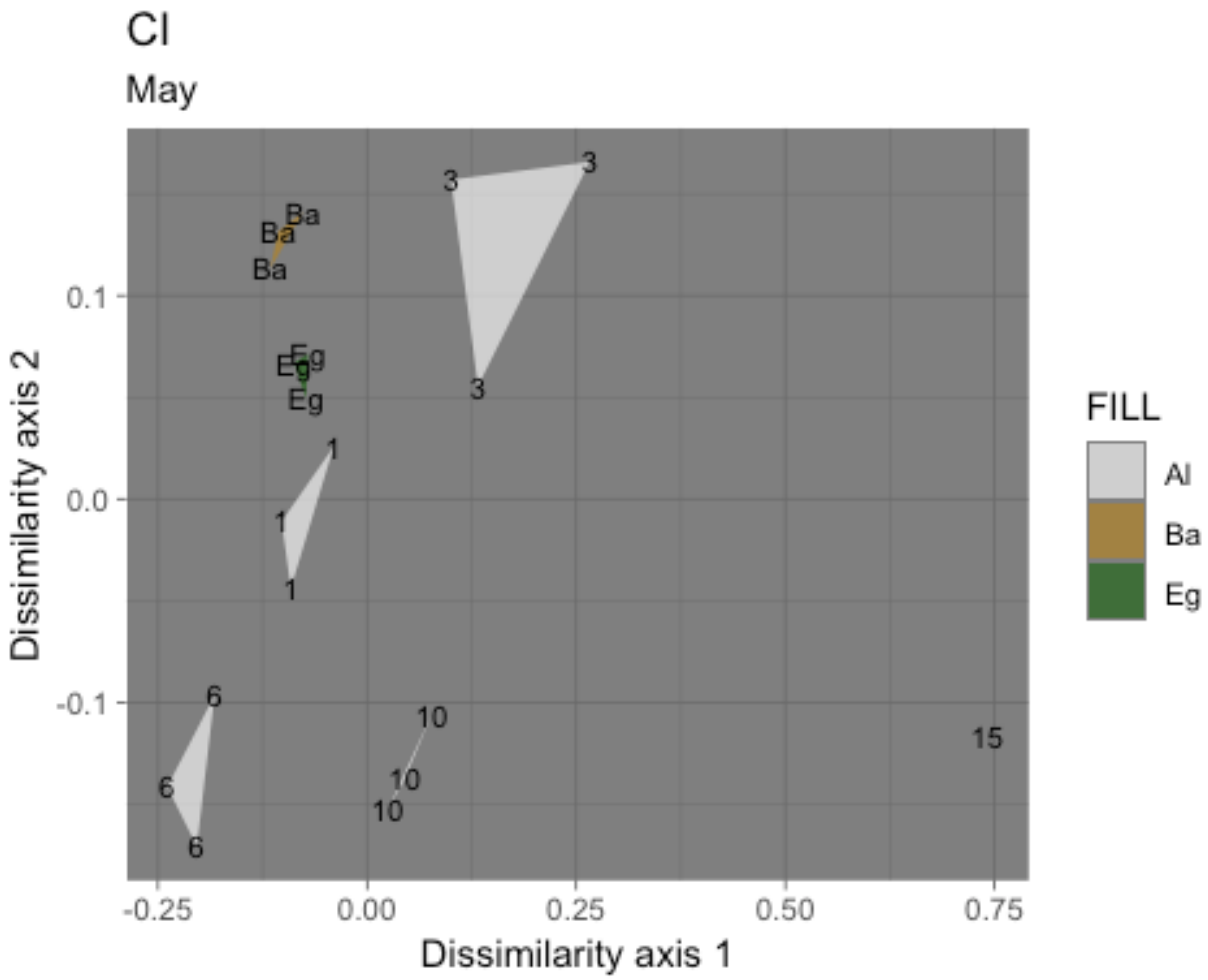


428

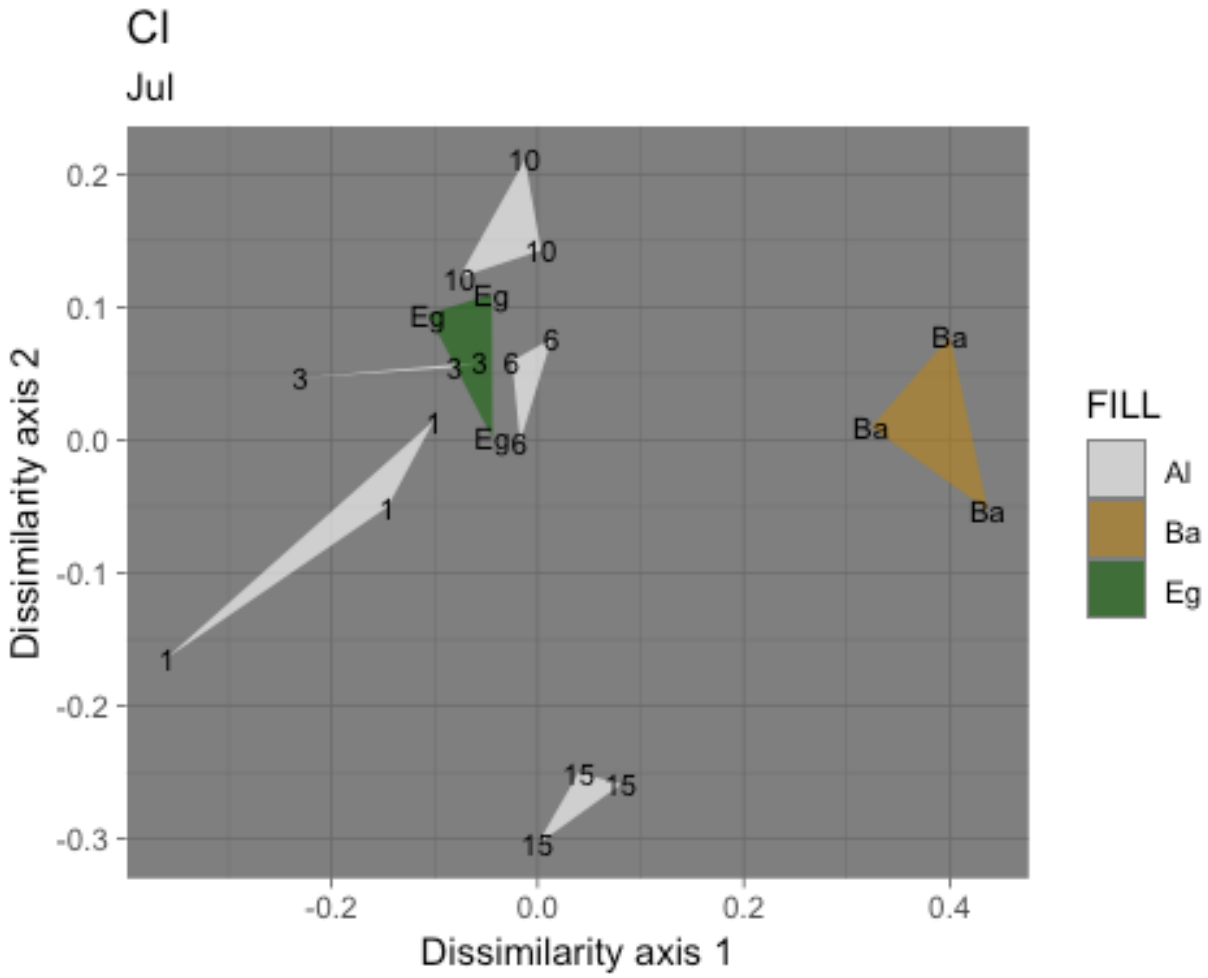
429 *Figure S6: Bray-Curtis dissimilarity between eDNA communities surveyed at adjacent points*

430 *along each full transect (all sites and months). Shading of violins indicates median spatial*

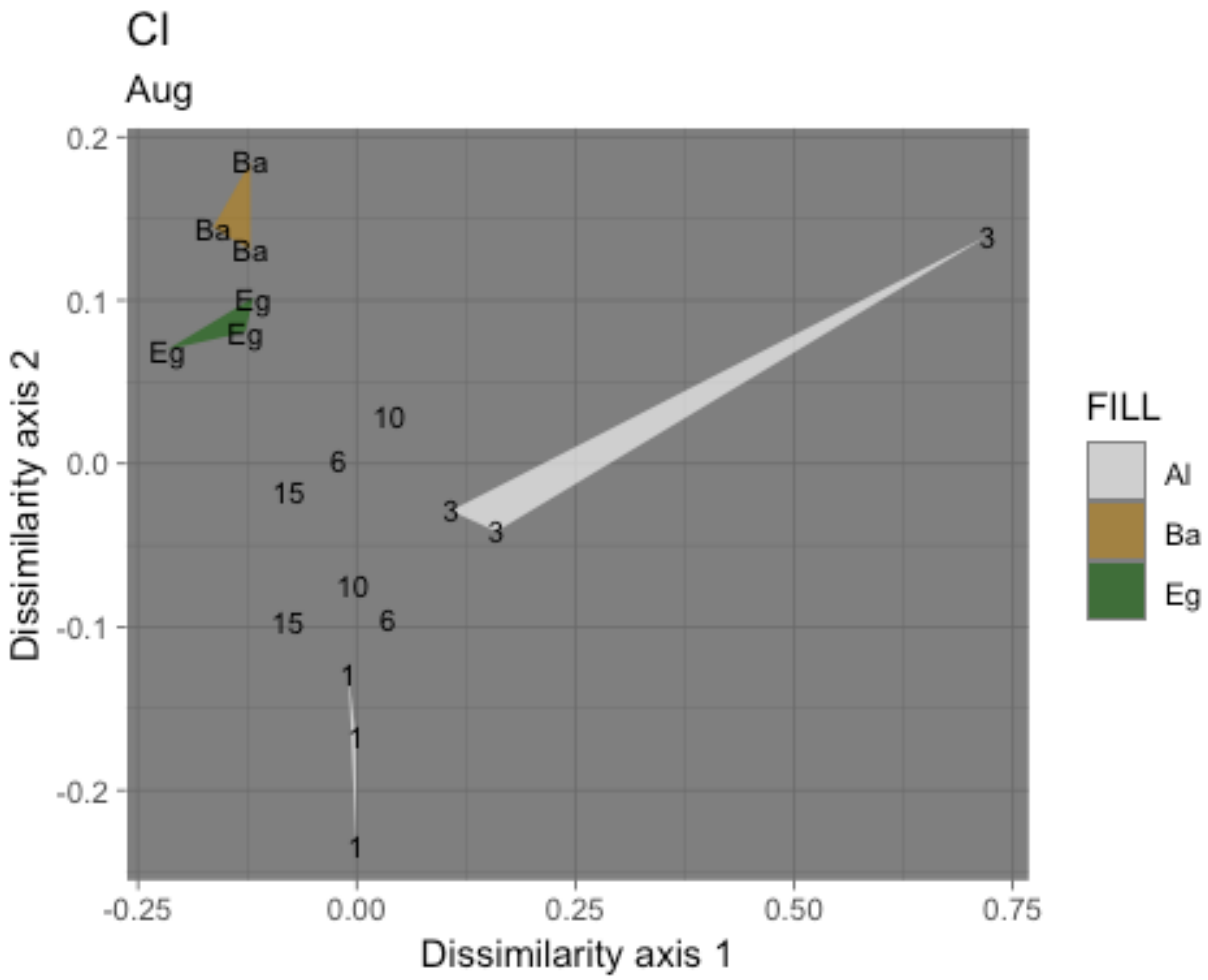
431 *distance between communities (dark = closer, light = more distant).*



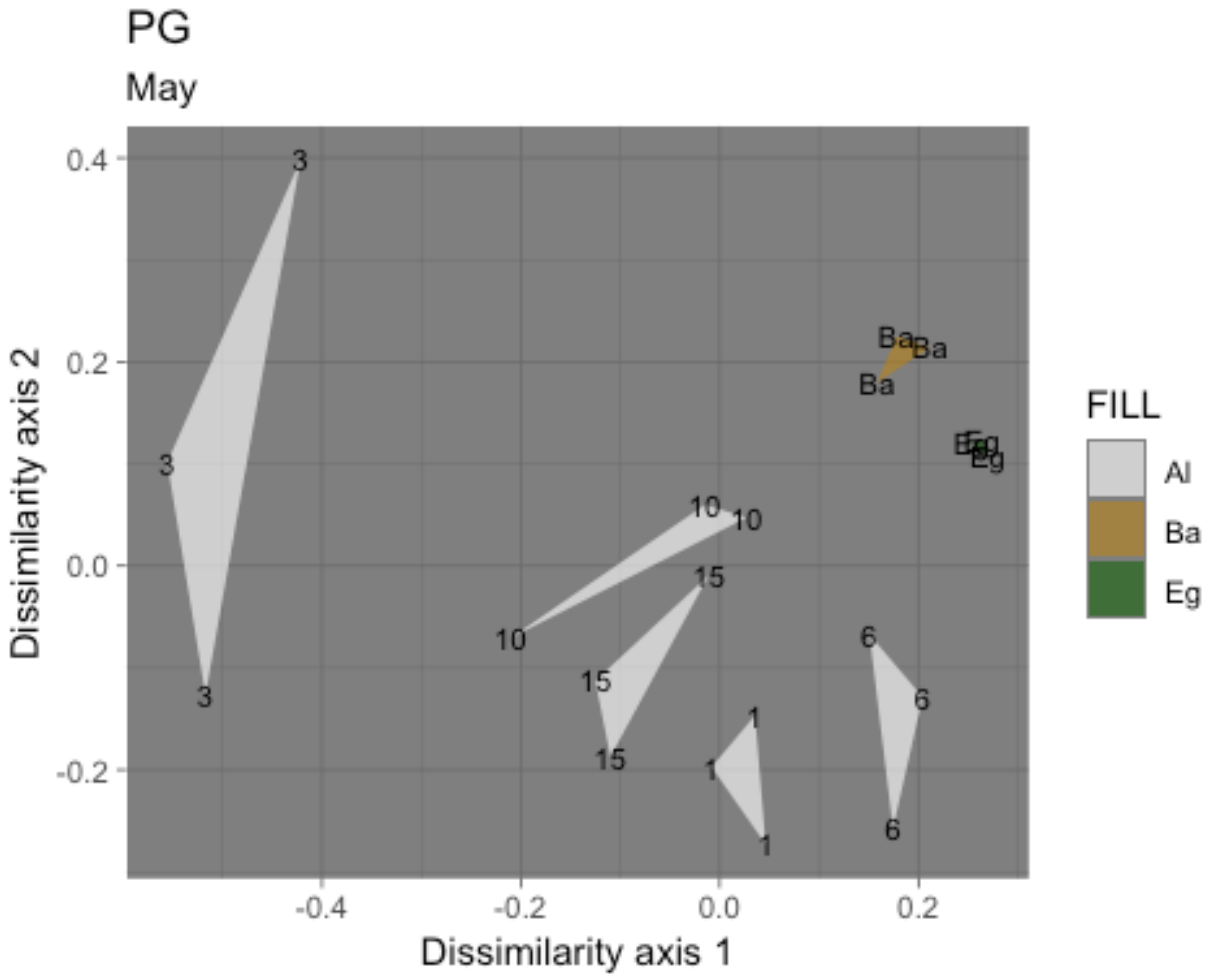
432



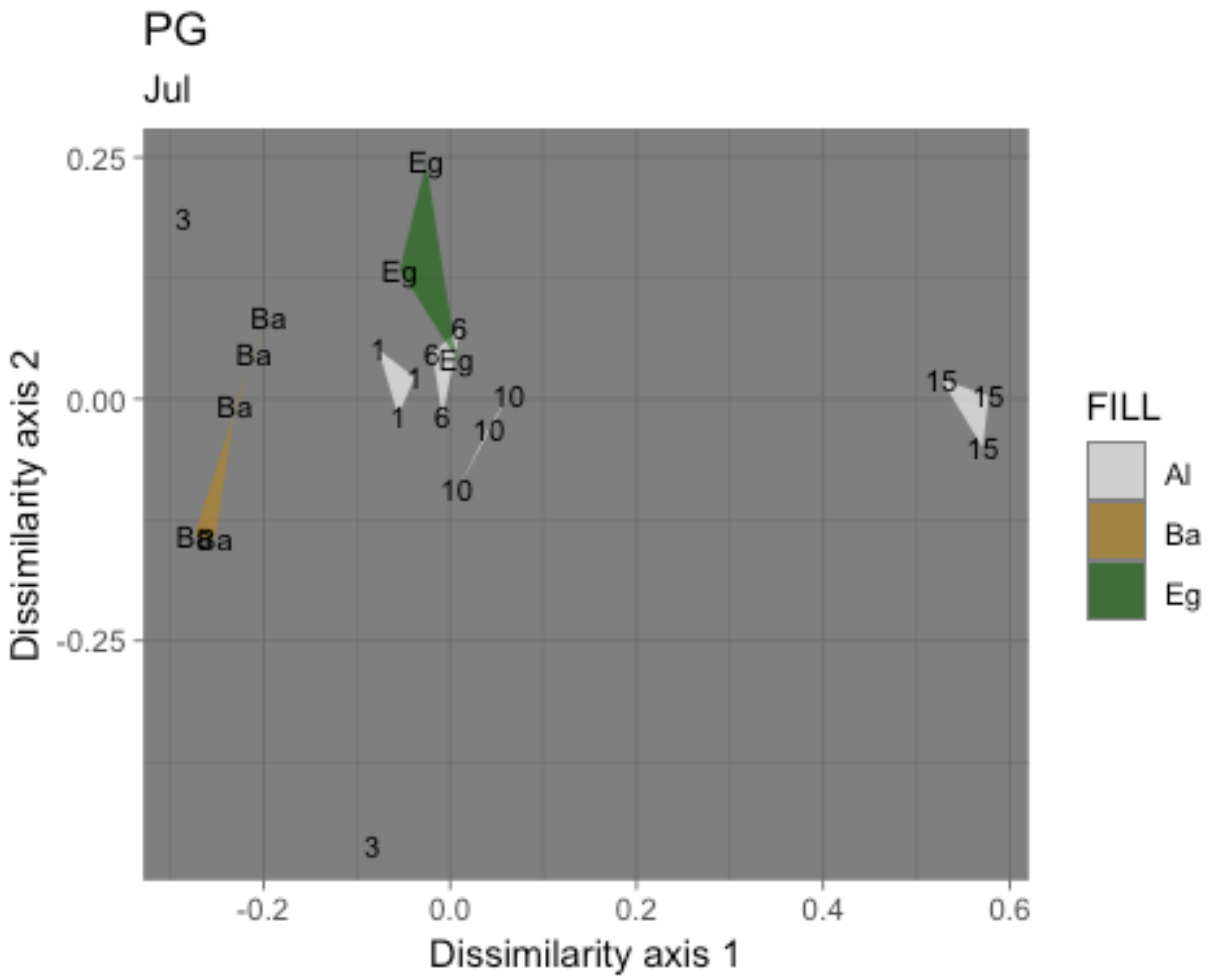
433



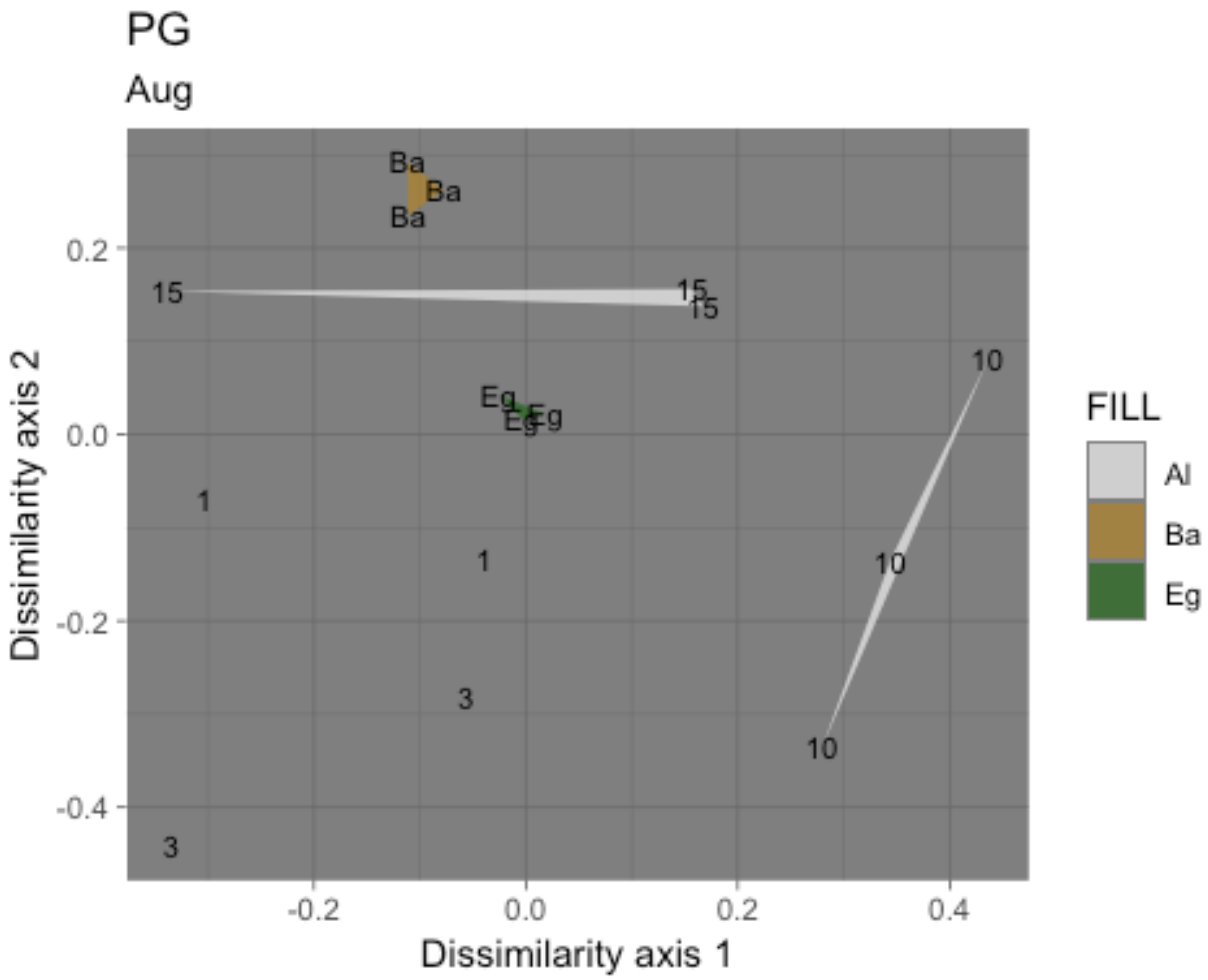
434



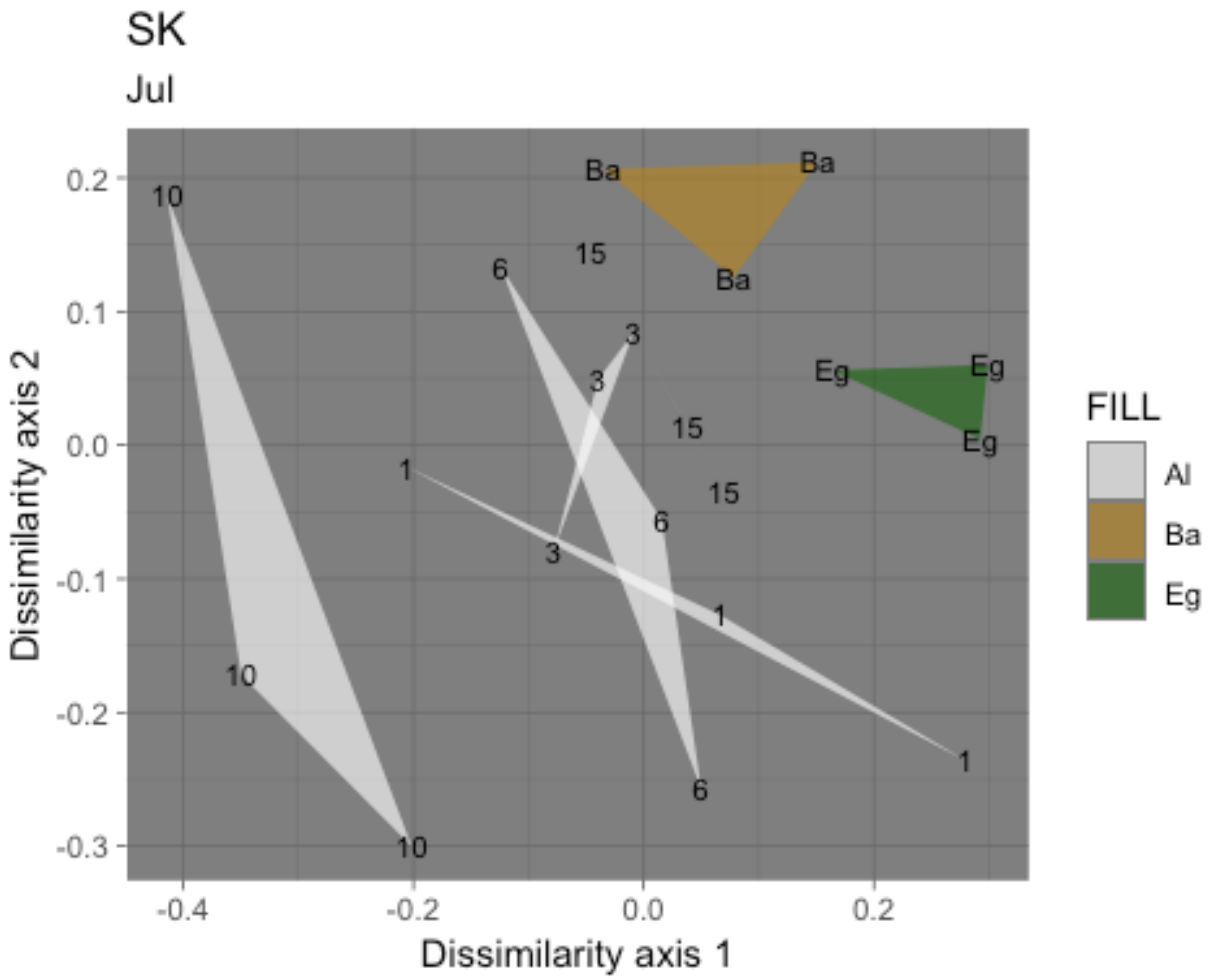
435



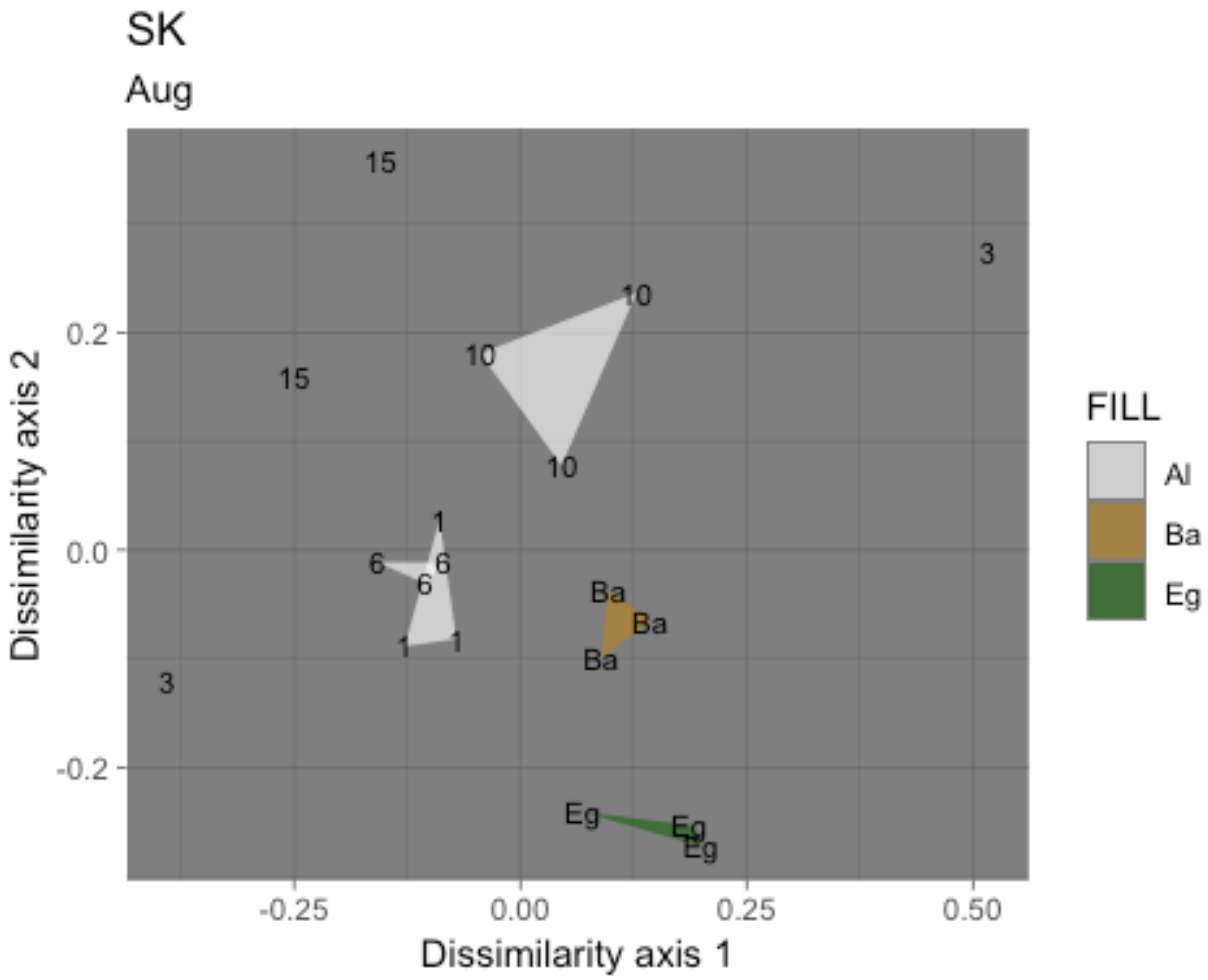
436



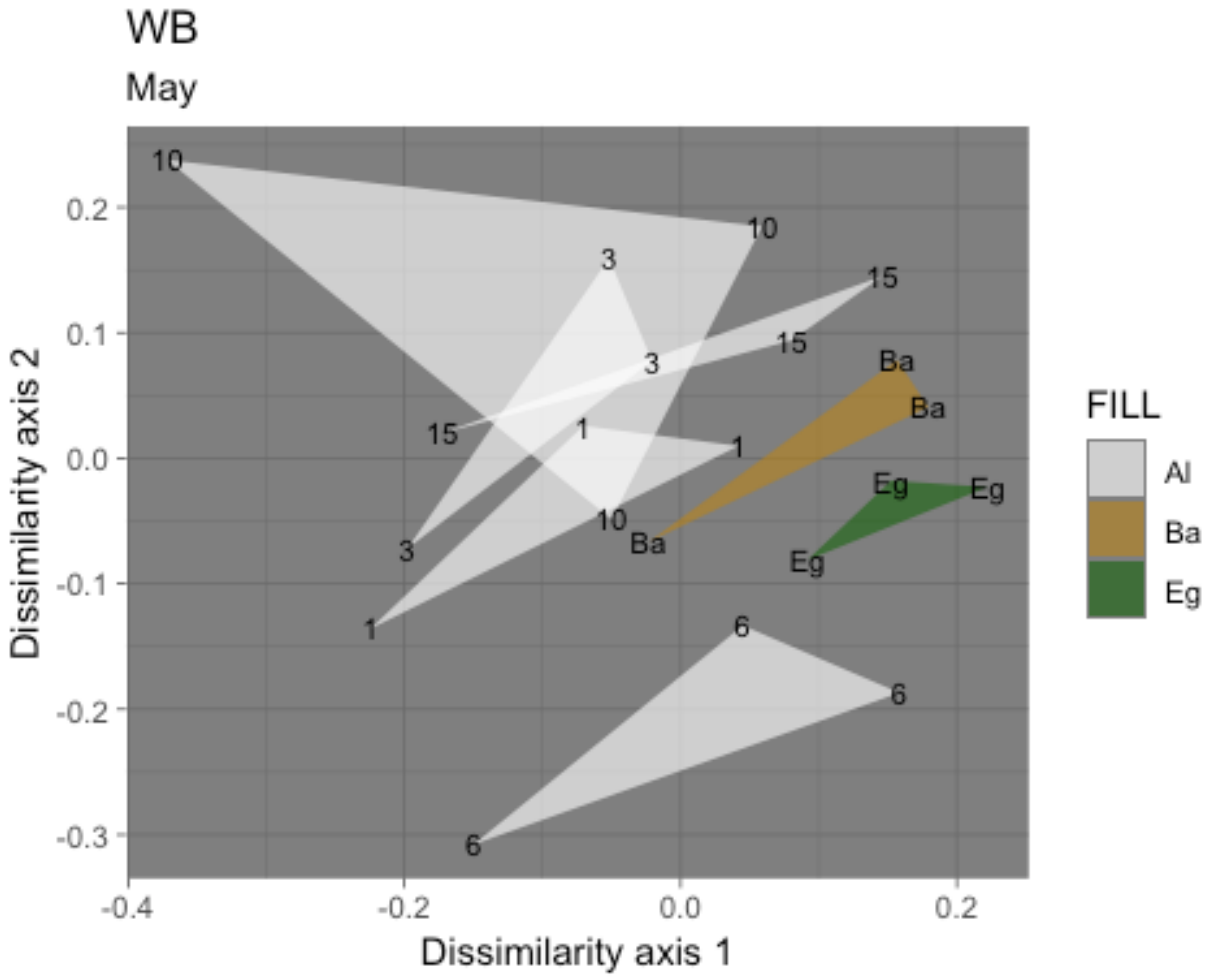
437



438



439



440

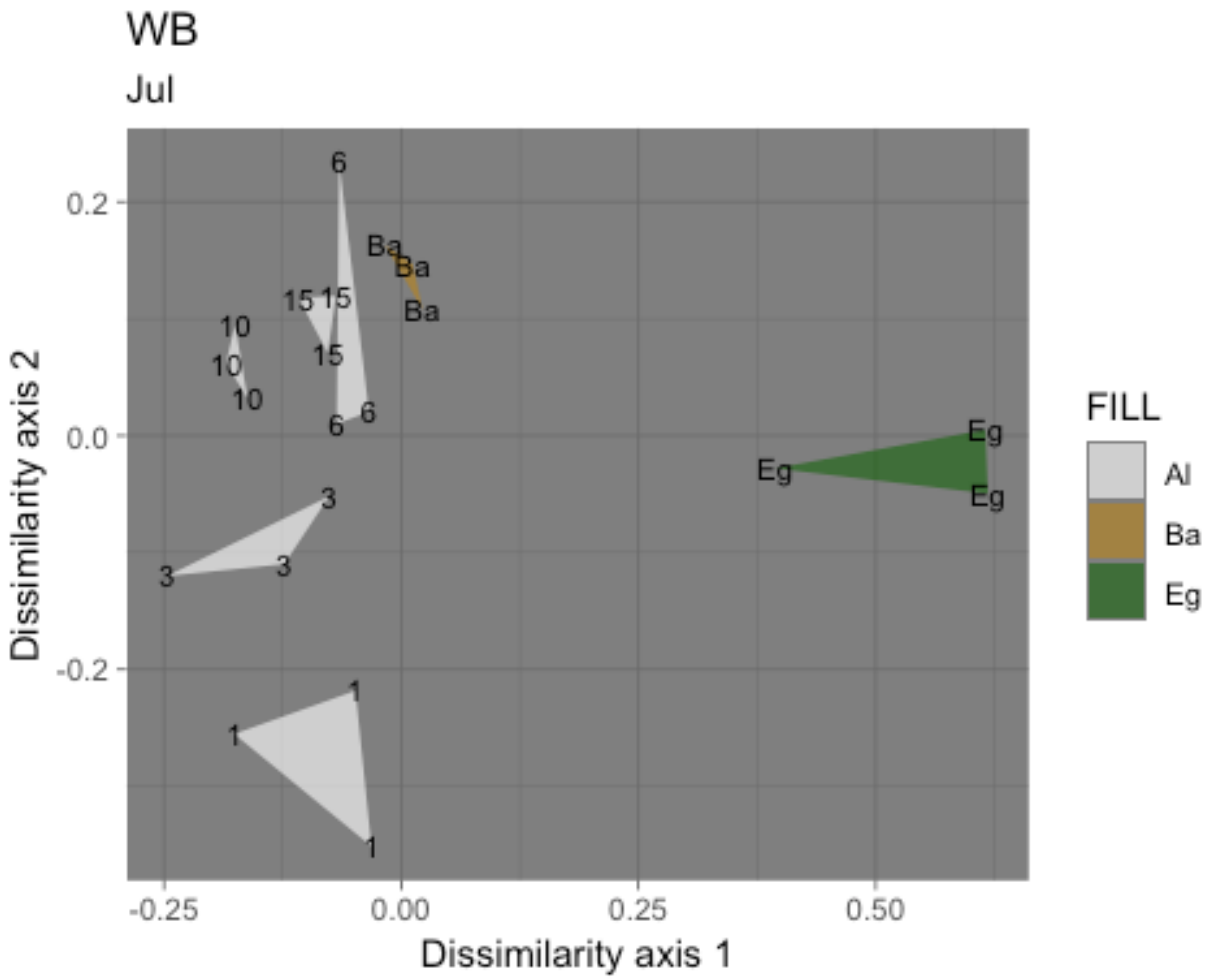


Figure S7: Ordination plots of samples along all ten fully-sampled transects from bare to eelgrass positions. Technical replicates of each biological sample are grouped as triangles. White alongshore transect samples are labeled with distance from eelgrass in meters; the single within-bed sample is green (labeled Eg) and the bare sample is brown (Ba).

References

Amelia Kolb Gabriela Hannach., Swanson L. 2016. *Marine phytoplankton monitoring program sampling and analysis plan*. Seattle, WA: King County Department of Natural Resources; Parks.

449 Callahan BJ., McMurdie PJ., Rosen MJ., Han AW., Johnson AJA., Holmes SP. 2016. DADA2:
450 High-resolution sample inference from illumina amplicon data. *Nature methods* 13:581.

451 Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., Madden TL. 2009.
452 BLAST+: Architecture and applications. *BMC Bioinformatics* 10:421.

453 Duffy JE. 2006. Biodiversity and the functioning of seagrass ecosystems. *Marine Ecology*
454 *Progress Series* 311:233–250.

455 Forrest BM., Keeley NB., Hopkins GA., Webb SC., Clement DM. 2009. Bivalve aquaculture in
456 estuaries: Review and synthesis of oyster cultivation effects. *Aquaculture* 298:1–15.

457 Fourqurean JW., Duarte CM., Kennedy H., Marbà N., Holmer M., Mateo MA., Apostolaki ET.,
458 Kendrick GA., Krause-Jensen D., McGlathery KJ., others. 2012. Seagrass ecosystems as a
459 globally significant carbon stock. *Nature geoscience* 5:505.

460 Gallego R.

461 Gibson J., Shokralla S., Porter TM., King I., Konynenburg S van., Janzen DH., Hallwachs W.,
462 Hajibabaei M. 2014. Simultaneous assessment of the macrobiome and microbiome in a bulk
463 sample of tropical arthropods through dna metasystematics. *Proceedings of the National*
464 *Academy of Sciences* 111:8007–8012.

465 Gobler CJ., Doherty OM., Hattenrath-Lehmann TK., Griffith AW., Kang Y., Litaker RW. 2017.
466 Ocean warming since 1982 has expanded the niche of toxic algal blooms in the north atlantic and
467 north pacific oceans. *Proceedings of the National Academy of Sciences* 114:4975–4980.

468 Groner ML., Burge CA., Cox R., Rivlin ND., Turner M., Van Alstyne KL., Wyllie-Echeverria
469 S., Bucci J., Staudigel P., Friedman CS. 2018. Oysters and eelgrass: Potential partners in a high
470 pCO 2 ocean. *Ecology* 99:1802–1814.

471 Gross EM. 2003. Allelopathy of aquatic autotrophs. *Critical reviews in plant sciences* 22:313–
472 339.

473 Harrison PG., Durance C. 1985. Reductions in photosynthetic carbon uptake in epiphytic
474 diatoms by water-soluble extracts of leaves of zostera marina. *Marine Biology* 90:117–119.

475 Heck Jr K., Hays G., Orth RJ. 2003. Critical evaluation of the nursery role hypothesis for
476 seagrass meadows. *Marine Ecology Progress Series* 253:123–136.

477 Hosack GR., Dumbauld BR., Ruesink JL., Armstrong DA. 2006. Habitat associations of
478 estuarine species: Comparisons of intertidal mudflat, seagrass (zostera marina), and oyster
479 (crassostrea gigas) habitats. *Estuaries and Coasts* 29:1150–1160.

480 Huson DH., Beier S., Flade I., Górska A., El-Hadidi M., Mitra S., Ruscheweyh H-J., Tappu R.
481 2016. MEGAN community edition-interactive exploration and analysis of large-scale
482 microbiome sequencing data. *PLoS computational biology* 12:e1004957.

483 Inaba N., Trainer VL., Onishi Y., Ishii K-I., Wyllie-Echeverria S., Imai I. 2017. Algicidal and
484 growth-inhibiting bacteria associated with seagrass and macroalgae beds in puget sound, wa, usa.
485 *Harmful algae* 62:136–147.

486 IOC Harmful Algal Bloom Programme and the World Register of Marine Species. Taxonomic
487 reference list of harmful micro algae.

488 Jones CG., Lawton JH., Shachak M. 1994. Organisms as ecosystem engineers. In: *Ecosystem*
489 *management*. Springer, 130–147.

490 Kelly RP., Gallego R., Jacobs-Palmer E. 2018. The effect of tides on nearshore environmental
491 dna. *PeerJ* 6:e4521.

492 Lahoz-Monfort JJ., Guillera-Arroita G., Tingley R. 2015. Statistical approaches to account for
493 false positive errors in environmental dna samples. *Molecular Ecology Resources*.

494 Lamb JB., Water JA van de., Bourne DG., Altier C., Hein MY., Fiorenza EA., Abu N., Jompa J.,
495 Harvell CD. 2017. Seagrass ecosystems reduce exposure to bacterial pathogens of humans,
496 fishes, and invertebrates. *Science* 355:731–733.

497 Leray M., Yang JY., Meyer CP., Mills SC., Agudelo N., Ranwez V., Boehm JT., Machida RJ.
498 2013. A new versatile primer set targeting a short fragment of the mitochondrial coi region for
499 metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents.
500 *Frontiers in zoology* 10:34.

501 Leray M., Knowlton N. 2015. DNA barcoding and metabarcoding of standardized samples
502 reveal patterns of marine benthic diversity. *Proceedings of the National Academy of Sciences*
503 112:2076–2081.

504 Li W., Cowley A., Uludag M., Gur T., McWilliam H., Squizzato S., Park YM., Buso N., Lopez
505 R. 2015. The embl-ebi bioinformatics web and programmatic tools framework. *Nucleic acids*
506 *research* 43:W580–W584.

507 Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads.
508 *EMBnet.journal* 17:pp–10.

509 Moore SK., Mantua NJ., Hickey BM., Trainer VL. 2009. Recent trends in paralytic shellfish
510 toxins in puget sound, relationships to climate, and capacity for prediction of toxic events.
511 *Harmful Algae* 8:463–477.

512 National Marine Fisheries Service West Coast Region. 2017. *WA eelgrass and shellfish*
513 *aquaculture workshop report.* Seattle, WA.

514 NOAA Fisheries. 2014. *California eelgrass mitigation policy and implementing guidelines*. West
515 Coast Region.

516 NOAA Fisheries. Habitat areas of particular concern.

517 Nordlund LM., Koch EW., Barbier EB., Creed JC. 2016. Seagrass ecosystem services and their
518 variability across genera and geographical regions. *PLoS One* 11:e0163091.

519 O'Donnell JL., Kelly RP., Lowell NC., Port JA. 2016. Indexed pcr primers induce template-
520 specific bias in large-scale dna sequencing studies. *PloS one* 11:e0148698.

521 Oksanen J., Blanchet FG., Kindt R., Legendre P., Minchin PR., O'hara R., Simpson GL.,
522 Solymos P., Stevens MHH., Wagner H., others. 2013. Package "vegan". *Community ecology*
523 *package, version* 2.

524 Orth RJ., Carruthers TJ., Dennison WC., Duarte CM., Fourqurean JW., Heck KL., Hughes AR.,
525 Kendrick GA., Kenworthy WJ., Olyarnik S., others. 2006. A global crisis for seagrass
526 ecosystems. *Bioscience* 56:987–996.

527 Port JA., O'Donnell JL., Romero-Maraccini OC., Leary PR., Litvin SY., Nickols KJ., Yamahara
528 KM., Kelly RP. 2016. Assessing vertebrate biodiversity in a kelp forest ecosystem using
529 environmental dna. *Molecular ecology* 25:527–541.

530 Potouroglou M., Bull JC., Krauss KW., Kennedy HA., Fusi M., Daffonchio D., Mangora MM.,
531 Githaiga MN., Diele K., Huxham M. 2017. Measuring the role of seagrasses in regulating
532 sediment surface elevation. *Scientific reports* 7:11917.

533 Puget Sound Partnership. Vital sign: Eelgrass.

534 R Core Team. 2016. *R: A language and environment for statistical computing*. Vienna, Austria:
535 R Foundation for Statistical Computing.

536 Renshaw MA., Olds BP., Jerde CL., McVeigh MM., Lodge DM. 2015. The room temperature
537 preservation of filtered environmental dna samples and assimilation into a phenol–chloroform–
538 isoamyl alcohol dna extraction. *Molecular ecology resources* 15:168–176.

539 Royle JA., Link WA. 2006. Generalized site occupancy models allowing for false positive and
540 false negative errors. *Ecology* 87:835–841.

541 Schnell IB., Bohmann K., Gilbert MTP. 2015. Tag jumps illuminated–reducing sequence-to-
542 sample misidentifications in metabarcoding studies. *Molecular Ecology Resources*.

543 Scrucca L., Fop M., Murphy TB., Raftery AE. 2016. Mclust 5: Clustering, classification and
544 density estimation using gaussian finite mixture models. *The R journal* 8:289.

545 Shelton AO., Francis TB., Feist BE., Williams GD., Lindquist A., Levin PS. 2017. Forty years of
546 seagrass population stability and resilience in an urbanizing estuary. *Journal of Ecology*
547 105:458–470.

548 Short F., Carruthers T., Dennison W., Waycott M. 2007. Global seagrass distribution and
549 diversity: A bioregional model. *Journal of Experimental Marine Biology and Ecology* 350:3–20.

550 Team RC., Worldwide C. 2015. The r stats package. *R Foundation for Statistical Computing:*
551 *Vienna, Austria*.

552 Trainer VL., Eberhart B-TL., Wekell JC., Adams NG., Hanson L., Cox F., Dowell J. 2003.
553 Paralytic shellfish toxins in puget sound, washington state. *Journal of Shellfish Research*
554 22:213–223.

555 Trainer V., Moore L., Bill B., Adams N., Harrington N., Borchert J., Silva D da., Eberhart B-T.
556 2013. Diarrhetic shellfish toxins and other lipophilic toxins of human health concern in
557 washington state. *Marine Drugs* 11:1815–1835.

558 Tubaro A., Dell'Ovo V., Sosa S., Florio C. 2010. Yessotoxins: A toxicological overview.
559 *Toxicon* 56:163–172.

560 Vera Trainer BB Teri King. 2016. *Soundtoxins manual*. Seattle, WA: NOAA/NMFS/NWFSC
561 Marine Biotoxins Program for Washington Sea Grant.

562 Washinton State Department of Ecology. 2013. *Fish consumption rates. Technical support*
563 *document: A review of data and information about fish consumption in washington.* Toxics
564 Cleanup Program, Olympia, WA: Department of Ecology, State of Washington.

565 Wiese M., D'agostino PM., Mihali TK., Moffitt MC., Neilan BA. 2010. Neurotoxic alkaloids:
566 Saxitoxin and its analogs. *Marine drugs* 8:2185–2211.