

SCIENTIFIC REPORTS



OPEN

Limited congruence exhibited across microbial, meiofaunal and macrofaunal benthic assemblages in a heterogeneous coastal environment

Sorcha Cronin-O'Reilly^{1,3,6}, Joe D. Taylor^{2,4}, Ian Jermyn¹, A. Louise Allcock¹, Michael Cunliffe^{2,5} & Mark P. Johnson¹

One of the most common approaches for investigating the ecology of spatially complex environments is to examine a single biotic assemblage present, such as macroinvertebrates. Underlying this approach are assumptions that sampled and unsampled taxa respond similarly to environmental gradients and exhibit congruence across different sites. These assumptions were tested for five benthic groups of various sizes (archaea, bacteria, microbial eukaryotes/protists, meiofauna and macrofauna) in Plymouth Sound, a harbour with many different pollution sources. Sediments varied in granulometry, hydrocarbon and trace metal concentrations. Following variable reduction, canonical correspondence analysis did not identify any associations between sediment characteristics and assemblage composition of archaea or macrofauna. In contrast, variation in bacteria was associated with granulometry, trace metal variations and bioturbation (e.g. community bioturbation potential). Protists varied with granulometry, hydrocarbon and trace metal predictors. Meiofaunal variation was associated with hydrocarbon and bioturbation predictors. Taxon turnover between sites varied with only three out of 10 group pairs showing congruence (meiofauna-protists, meiofauna-macrofauna and protists-macrofauna). While our results support using eukaryotic taxa as proxies for others, the lack of congruence suggests caution should be applied to inferring wider indicator or functional interpretations from studies of a single biotic assemblage.

Taxa are typically thought to vary in their tolerance of environmental stresses and contaminants^{1,2}. This is the basis of biotic indices like the AZTI Marine Biotic Index (AMBI) used to summarize ecological quality from the relative abundance of macroinvertebrates classified into groups representing different tolerance levels². For marine sedimentary environments, measurements of ecological quality often involve macrofauna³. Other groups, such as bacteria, protists or archaea, are surveyed less frequently⁴⁻⁶. It is not clear to what extent the response of different groups to environmental gradients is congruent, i.e. respond in similar ways. If distributions of different groups lack congruence, the generality of indices derived from a single assemblage are restricted. Without congruence, more consideration must be given to choosing the most suitable biotic assemblage for the monitoring purpose and context (i.e. macrofauna vs. meiofauna⁷ or bacteria⁸, oil rig vs. coastal monitoring). In addition, variability in ecosystem function is unlikely to be predictable from a single assemblage⁹ if other functionally important groups respond in different ways to shared environmental conditions.

¹Ryan Institute & School of Natural Sciences, NUI Galway, University Road, Galway, H91 TK33, Ireland. ²Marine Biological Association of the UK, The Laboratory, Citadel Hill, Plymouth, PL1 2PB, UK. ³Centre for Sustainable Aquatic Ecosystems, Harry Butler Institute, Murdoch University, 90 South St, Murdoch, 6150, Western Australia, Australia. ⁴University of Salford, School of Environment & Life Sciences, Peel Building, Salford, M5 4WT, UK. ⁵Marine Biology and Ecology Research Centre, School of Biological and Marine Sciences, Plymouth University, Drake Circus, Plymouth, PL4 8AA, UK. ⁶School of Veterinary and Life Sciences, Murdoch University, 90 South St, Murdoch, 6150, Western Australia, Australia. Correspondence and requests for materials should be addressed to S.C.-O. (email: sorcha.cronin-o'reilly@murdoch.edu.au)

Received: 9 May 2018

Accepted: 4 October 2018

Published online: 19 October 2018

A lack of congruence has been observed previously, among sedimentary taxa^{5,10}. Benthic archaea and bacteria have responded separately to shared environmental variables in the North Sea⁵ while nematode and microbial assemblages have shown non-congruent patterns in heavy-metal contaminated soils from an explosive manufacturing plant in Scotland¹⁰. However, some spatial variation in benthic assemblages has been found to be congruent across groups, for example between macrofauna and foraminifera around a sewage sludge disposal site in the Firth of Clyde⁶ or nematodes and macrofauna at a coastal site in Brazil¹¹. Studies of similarity in response have often examined a clearly identified gradient of impact associated with oil spills or fish farm cages^{12,13} and also include manipulated increases in temperature¹⁴. Levels of contaminants and environmental conditions may, however, covary in complex ways. For example, different contaminants can come from spatially separated sources. In coastal environments, natural environmental conditions will vary with factors such as depth and wave exposure. In addition, anthropogenic contaminants like hydrocarbons and metals are commonplace stressors in these areas. This sort of multiple-stress context provides a more demanding test of congruence in taxon responses to environmental and anthropogenic stresses.

An example of a potentially complex seafloor with various sources of contaminants is provided by Plymouth Sound (Southwest England). The City of Plymouth has around 250,000 inhabitants and is home to a long-established naval dockyard. The Tamar and Plym rivers drain into the Sound and the catchment contains now-abandoned metal mines and there are reports of relatively high hydrocarbon levels¹⁵. We examined the congruence of five benthic groups (archaea, bacteria, microbial eukaryotes/protists, meiofauna and macrofauna) to multiple gradients in Plymouth Sound. Each group contributes to a series of key ecosystem services such as nutrient cycling, primary production and trophic transfer of energy^{16,17}, but rarely are they considered together. Potential explanatory variables were sediment granulometry, hydrocarbon content, metal concentrations, nutrients and organic matter. The null hypothesis was that, in an environment with multiple gradients, there will be no congruence in the relationships among sites as expressed by the taxa present in different groups. As macrofauna can influence other groups by sediment reworking, an index of bioturbation (community bioturbation potential, BP_c)¹⁸ and the distribution of different functional types of bioturbation were derived to test whether these variables influenced the occurrence of taxa within groups.

Results

Environmental and contamination gradients. Principal components analysis (PCA) identified a number of different gradients across the sampled sites in Plymouth Sound (Fig. 1). PCA scores from different sets of variables were generally not correlated, implying that the data reduction process identified independent gradients across samples (mean pairwise correlation 0.02, no PCA score pairs significantly correlated). The main axis for granulometry variables separated sites across a spectrum of particle sizes (PC1 associated with 76% of the variation between sites, Supplementary Table 1). Cawsand Bay was the site with the highest sand particle fraction, with Sutton Lock dominated by silts. Hydrocarbons were generally positively correlated, resulting in a first principal component sorting sites by overall contamination level (95% of variance associated with PC1 hydrocarbon, Supplementary Table 2). The PCA of environmental variables contained a gradient stressing covariation in arsenic and a number of metals (Co, Cr, Cd, Cu, Zn and Pb). This identified Cawsand Bay as a less contaminated site and Sutton Lock as the most heavily contaminated. Surficial modifier and biodiffusor functional types were positively correlated and more associated with the coarser sediment sites, with upward/downward conveyors more common at West Mud.

Phylogenetic and macrofaunal diversity. A total of 5,566 operational taxonomic units (OTU) were identified to various taxonomic levels as active constituents across archaea, bacteria, protists and meiofauna (1543, 2126, 1791 and 106 OTUs respectively) and total number of species/OTUs was calculated for each group (Supplementary Table 3). Broadly, archaeal rRNA transcript sequence libraries were characterised by representatives from three major phyla; Crenarchaeota, Euryarchaeota and Parvarchaeota. The novel uncharacterised archaeal order YLA114 was frequently detected in sequence libraries at all sites (relative abundance of 46–82%) except from Sutton Lock, where the rRNA transcript library was dominated by the Nitrosopumilales OTUs (74%). Bacterial rRNA transcript sequence libraries across all sites were dominated by the phyla Bacteroidetes, Firmicutes and Proteobacteria. Order Clostridiales (Firmicutes) was relatively abundant at Cawsand Bay (45%) and rare at all other sites (2–4%). Desulfobacteriales were largely present at all sites, relatively highest at Inner Breakwater (21%) and lowest at Cawsand Bay (8%). The Cyanobacteria order Crocococcales was relatively abundant in the sequence libraries for Sutton Lock (15%) compared to all other sites (0–0.75%).

Most of the protists in the 18S rRNA transcript sequence libraries were Ochrophyta, specifically diatoms, except at Mallard Shoal where Ciliophora had a higher relative abundance (43%) in the rRNA transcript library compared to all other sites (0–8%). Foraminifera, class Monothalamea, were most abundant in the rRNA transcript library for Sutton Lock. Platyhelminthes and Nematoda made up a large fraction of meiofaunal reads in 18S rRNA transcript sequence libraries. Sequence reads for the Platyhelminthes were most abundant at St. John's Lake, while Nematoda were dominant in sequence libraries at Jenyncliff Bay (70%, 66% respectively). Nematoda and Copepoda are often classed as meiofauna. Our macrofaunal definition is consistent with standard, sieve-based benthic survey methods (organisms larger than 500 µm¹⁹). Both of these taxa were therefore recorded as part of the macrofaunal assemblage, with smaller nematodes in the meiofauna. The macrofaunal assemblage included 216 taxa mostly identified across two phyla, Annelida and Arthropoda (Supplementary Table 4). The polychaete family Cirratulidae were most abundant at Sutton lock and Inner Breakwater, but all other sites had Cossuridae (detritus feeding polychaetes) or Ampharetidae (bristle worms) as the most abundant taxa. The highest abundance of Arthropoda was found at Sutton Lock (19%) due to a large presence of Aoridae (amphipod) females. Pollution tolerant taxa (e.g. *Aphelocheata marioni*) were more common at Sutton Lock.

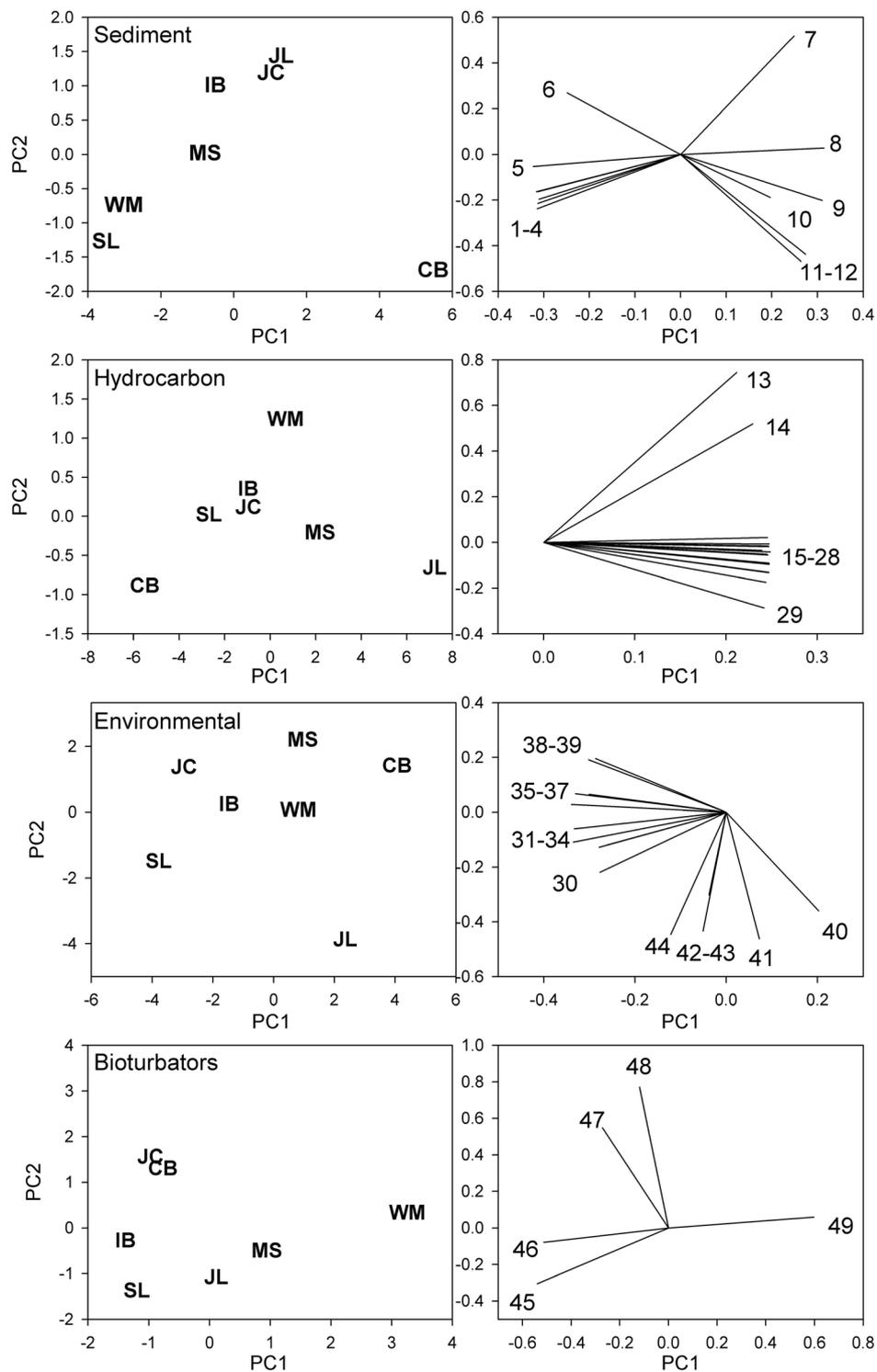


Figure 1. Principal components analysis of sediment, hydrocarbon, environmental and bioturbator variables across sites. Left hand panels show site scores used as predictor variables in subsequent canonical correspondence analysis. Right hand panels show loadings of variables on the first and second principal components. Site labels; CB) Cawsand Bay, IB) Inner Breakwater, JC) Jennycliff Bay, MS) Mallard Shoal, WM) West Mud, JL) St. John's Lake, SL) Sutton Lock. Variable labels; 1) % medium silt, 2) fine silt, 3) coarse silt, 4) very fine silt, 5) clay, 6) very coarse silt, 7) very fine sand, 8) fine sand, 9) medium sand, 10) coarse sand, 11) very coarse sand, 12) very fine gravel, 13) acenanthrene, 14) phenanthrene, 15) fluorene, 16) pyrene, 17) benzo(b)fluoranthene, 18) benzo(g,h,i) perylene, 19) chrysene, 20) anthracene, 21) Total PAHs, 22) acenaphthylene, 23) benzo(k)fluoranthene, 24) indeno(1,2,3-cd)pyrene, 25) benzo(a)pyrene, 26) benz(a)anthracene, 27) dibenzo(a,h)anthracene, 28) fluoranthene, 29) naphthalene, 30) total N, 31) total P, 32) Zn, 33) Pb, 34) Cu, 35) Cd, 36) As, 37) S (SO_4), 38) Cr, 39) Co, 40) C (CaCO_3), 41) total C, 42) C (organic), 43) N (NO_3), 44) Hg, 45) surficial modifiers, 46) biodiffusers, 47) downward conveyors, 48) upward conveyors and 49) upward/downward conveyors.

The resulting Infaunal Quality Index (IQI)³ value for Sutton Lock was 0.83, as it was at St. Johns Lake. An ecosystem health assessment based on macrofauna would therefore rank Sutton Lock and St John's Lake as the most impacted sites, although all sites lie above the standard threshold (0.75) used to describe ecosystem status as high (Supplementary Table 5).

Congruence in assemblage structure. Results from canonical correspondence analysis (CCA)²⁰ differed across taxonomic groups with varying numbers of predictor variables significantly associated with inter-site variation in assemblage structure (Fig. 2). Archaeal and macrofaunal assemblages did not vary in response to gradients of any of the predictor variables examined. In contrast, sediment, environmental and bioturbation-related variables were associated with variation in the bacterial assemblage. Variation in both protist and meiofauna was related to hydrocarbon gradients, with additional roles for sediment structure and environmental variables in the former and bioturbation in the latter.

Although there appear to be some consistently outlying sites in the CCA plots, these patterns were not universal. Archaea, bacteria and macrofauna at Cawsand Bay were clearly separated from other sites, but the same site was not an outlier for meiofauna or protists. Out of 10 possible comparisons, only the PCO ordination pairs of meiofauna-protists, meiofauna-macrofauna and protist-macrofauna were congruent ($p < 0.05$, PROTEST randomisation tests). The average correlation between dissimilarity matrices for these congruent pairs was 0.631 (SE 0.097) compared to 0.141 (SE 0.098) for the other seven pairs.

Discussion

The variations in assemblage structure across sites were not generally congruent across benthic groups: only three out of the 10 possible group pairs showed congruence in inter-site dissimilarities with congruence only evident between eukaryotic groups. Previous studies have investigated congruence among eukaryotes in the spatial patterns of macrofauna, meiofauna and protists. Nematode and macrofaunal samples have been seen to have congruent patterns of assemblage structure in Araçá Bay, Brazil¹¹. The environmental gradients associated with glacial meltwater in a fjord were associated with congruence of foraminifera and macrofauna²¹.

A lack of congruence has been shown, however, when comparisons spread beyond eukaryotic groups. Similarities in response among benthic taxa were investigated between seven benthic sites in the North Sea and none of the groups investigated (bacteria, archaea, ammonia-oxidising bacteria and ammonia-oxidizing archaea) were congruent⁵. In a study of contaminated soils, relationships were found between soil chemistry and PCA-reduced microbial variables (such as plate counts, proportion of metal resistant bacteria and marker fatty acids), but not between univariate summaries of nematode diversity and the soil or microbial variables¹⁰. The current study is the first study to investigate all five major benthic groups simultaneously and it further supports an absence of congruence between eukaryotes and prokaryotes, or between archaea and bacteria alike.

The patterns observed in Plymouth Sound do not reflect separation into concordant and non-concordant groups by sampling or identification methods (core samples *vs.* subsample, morphological *vs.* molecular). Molecular *vs.* morphological approaches, although kept constant in some studies showing congruence^{6,11}, are also constant in others where congruence is lacking⁵. Congruence in Plymouth Sound is evident between a morphologically identified group (macrofauna) and two molecularly defined groups (protists and meiofauna) whereas it is lacking between some of the molecularly defined pairs. This suggests that the results are not determined by differences in the methods used for group characterisation.

There are clearly differences between the evidence from molecular methods and direct counts. An important issue for measures of diversity and richness is that, when sequencing rRNA genes or transcripts from eukaryotic organisms, relative abundances of sequence reads often don't relate directly to individuals. This is due to variation in rRNA copy numbers per cell/individual. Microeukaryotes such as ciliates may have many thousands of copies of rRNA genes within a single cell²² and with macrofauna and meiofauna, the size of the organism will also influence representation in sequence libraries²³. Therefore, statistical tests using presence/absence data are advisable, as we have used here when examining congruence (PROTEST and Mantel tests). As RNA was used here, this could create a bias towards representation of the more active members of the community, reflecting higher transcript copy numbers in actively metabolizing cells²⁴. In sedimentary systems, using RNA may be important to overcome the potential problem of relic DNA in sediment^{25,26} or the detection of inactive or resting cells that have settled from the water column onto the sediment. However, many studies have used RNA and DNA in assessing sedimentary microbial communities and found results to be largely comparable in terms of species richness^{24,27}. Despite differences in methodology, several studies have shown that molecular techniques are comparable to direct counts for aquatic monitoring and ecological patterns are the same across pollution gradients^{28,29}. In the assessment of meiofauna and macrofauna, it is important to note that while the 18S rRNA gene is wholly suitable for broad assessment of richness of taxa within benthic systems, the mitochondrial marker Cytochrome c oxidase I (COI) will give more accurate species-level taxonomic assignments, although this marker is limited by very few reference database sequences in comparison to the 18S rRNA gene³⁰.

Molecular techniques are becoming more widely used in aquatic monitoring and biodiversity assessments and several detailed reviews have now been produced outlining the methods and limitations^{31–34}. However, laboratory methods and bioinformatics analyses remain highly variable between studies, making cross comparison at the OTU level difficult. For example, different DNA and RNA extraction methodologies extract different taxonomic groups more efficiently, particularly for bacteria³⁵, leading to sometimes very different patterns of OTU richness. Also, different bioinformatic pipelines can influence OTU richness³⁶. The approaches we have taken in this study have been used across a number of other studies, but ultimately the community needs to develop standardized approaches for aquatic monitoring using molecular techniques.

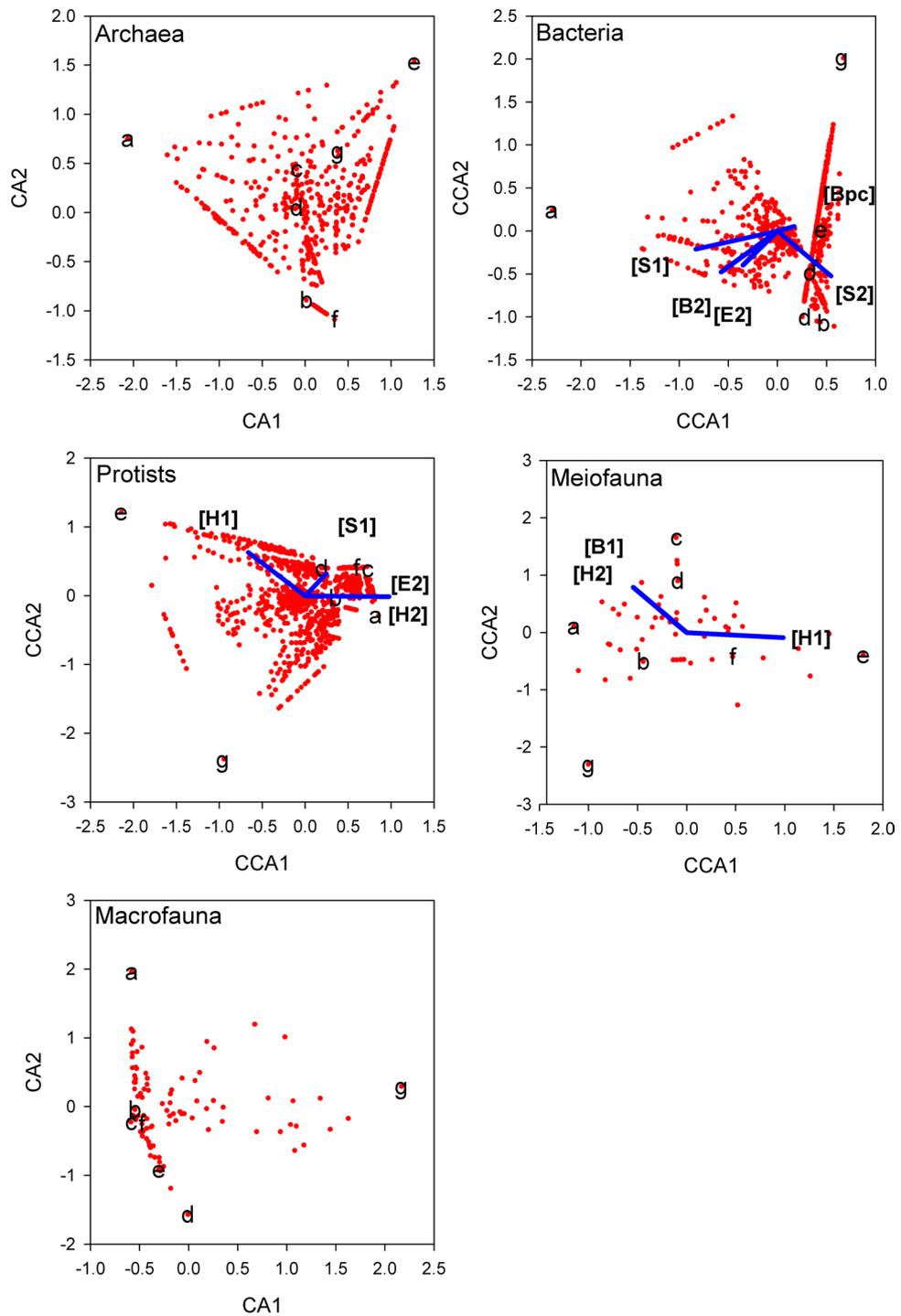


Figure 2. Correspondence analysis plot for each taxon group. Red points are the positions of individual taxa with respect to sites. Where predictor variables were significantly associated with pattern across the sites, these are added to the plots as lines, with direction and magnitude from the origin indicating the influence of the variable. Site labels; (a) Cawsand Bay, (b) Inner Breakwater, (c) Jennycliff Bay, (d) Mallard Shoal, (e) West Mud, (f) St. John's Lake, (g) Sutton Lock. Variable labels; S1) sediment PC score on axis 1, S2) sediment PC2, H1) hydrocarbon PC1, H2) hydrocarbon PC2, E2) environmental PC2, B1) bioturbation PC1, B2) bioturbation PC2 and BP_c) community bioturbation potential index.

There may be size-related aspects of how taxa respond to the perceived environmental gradients³⁷, but these are not easily identified given the size differences between macrofauna and other groups and the lack of congruence between similar-sized taxa (i.e. archaea and bacteria). A lack of congruence may also reflect the divergence of stress responses in groups with different evolutionary histories³⁸. In contrast, recent findings support similar

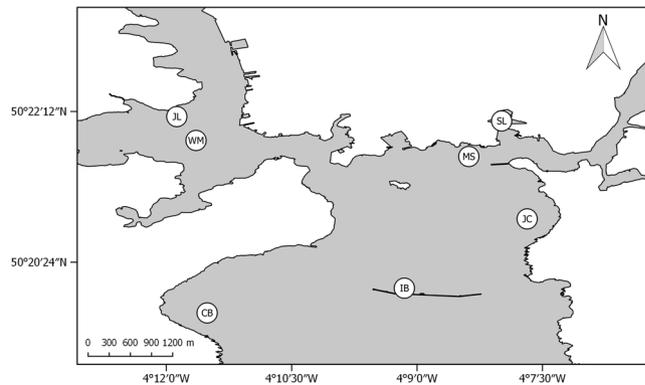


Figure 3. Plymouth Sound site map. Site labels; CB) Cawsand Bay, IB) Inner Breakwater, JC) Jennycliff Bay, MS) Mallard Shoal, WM) West Mud, JL) St. John's Lake, SL) Sutton Lock. Map created using QGIS version 2.14 [QGIS Development Team (2016). QGIS Geographic Information System. <http://qgis.org> .

responses to temperature increase across microbes and metazoans, irrespective of taxon size and phylogeny¹⁴. This may indicate that, when considering a dominant changing factor such as temperature, congruence may occur across divergent lineages.

Of course, the results from Plymouth Sound may be specific to the extent and type of gradients present. As anticipated, Plymouth Sound sediments showed a number of independent gradients of environmental contaminants. Responses to these gradients were identified for three (bacteria, protists and meiofauna) of the five groups studied and different combinations of variables were associated with changes in assemblage structure for each case. Congruence may emerge in situations where the environmental gradients are more aligned (e.g. synchronous increase/decrease of variables) and the range of each gradient is greater (e.g. around well-defined pollution sources) or where the spatial scale of the study is larger such that smoothing of small-scale variability with respect to larger regional patterns occurs³⁹. Additional limitations may exist with the temporal setting of this study, as it represents one time point. The season of sampling may have increased the abundances of particular taxa, e.g. Cyanobacteria and the uncharacterised order YLA114 recognized to dominate during autumn⁴⁰. It is unclear whether these patterns will remain constant through seasons as most assemblage structures vary temporally; for example through recruitment events linked to seasonal fluctuations in resources⁴¹.

Beyond the potential issues in assuming one group can act as a proxy for another, a lack of congruence has implications for interpreting the ecosystem function and health of benthic sediments on the basis of sampling a single assemblage. Based on rRNA sequencing of groups, the results show active groups responding to separate environmental gradients while displaying different degrees of taxon turnover between sites. As all groups impact ecosystem function, this therefore implies that interpretations of ecosystem function based on one group will potentially be compromised by patterns in another⁴². As of yet, any differences in archaea, bacteria, protist and meiofauna among sites are difficult to interpret as biodiversity ecosystem function (BEF) relationships are little studied for these groups⁴³. It seems likely that differing patterns of biodiversity across taxa may add to the potential causes of context-dependency already identified in BEF relationships^{18,44}.

Restrictions on generalization caused by a lack of congruence extend further when considering ecosystem monitoring through biotic indices (e.g. AMBI)². The congruence between spatial turnover in taxa for meiofauna, macrofauna and protists supports the application of these groups as proxies for each other, provided the temporal setting of the monitoring programmes remains constant. It may be advisable, however, to include an index of the archaea/bacteria consortium, if such an index is available for the ecosystem (e.g. MC-IBI for rivers)¹. The result of eukaryotic congruence extends to a functional index (BP_c) derived from macrofauna¹⁸, which was associated with variation in bacterial composition. Alternative variables derived from the composition of bioturbator functional types were also associated with variation in bacteria and meiofauna. While these results emphasise that trait-based approaches can potentially define the links between taxa (including eukaryote/prokaryote interplay), the approach will need further research to establish the repeatability of such associations and the mechanisms of interaction.

Material and Methods

Study area and sediment sampling. Benthic samples were taken using a HAPS corer over two days (12th-13th of September 2013) at seven sites in Plymouth Sound (Fig. 3); Cawsand Bay (CB), Inner Breakwater (IB), Jennycliff Bay (JC), Mallard Shoal (MS), West Mud (WM), St. John's Lake (JL) and Sutton Lock (SL). Three cores (15 cm diameter, 30 cm depth) were collected for measurement of environmental variables, microbial (i.e. archaea, bacteria and protists) and meiofaunal diversity analysis. Microbial and meiofaunal community analyses were based on three randomly spaced small cores (1 cm diameter, 5 cm depth) from each larger core and samples were placed into 1.5 mL centrifuge tubes and frozen immediately on dry ice for transfer to the laboratory where they were stored at -80°C . A single replicate sample from each site was used for molecular analyses of these assemblages. A 200 g subsample was sliced from the top 5 cm layer of each larger core for hydrocarbon fingerprinting, trace metal and nutrient analysis. A single 100 g subsample from each site was used for granulometry

and organic carbon content. Subsamples were stored in Rilsan® bags at -20°C . Three additional cores (15 cm diameter, 30 cm depth) were taken for morphological characterisation of the macrofaunal community.

Laser particle sizing of sediments used nine 5 mg pseudo-samples of homogenised sediment samples from each site. Grain size distribution summary statistics were generated using GRADISTAT⁴⁵. Organic content of samples were estimated using loss on ignition (LOI) where wet sediment samples were oven dried (24 hr at 100°C) before being ignited in a muffle furnace (6 hr at 450°C). The percentage weight loss after ignition was taken as a measure of total organic matter. Aliphatic and aromatic hydrocarbons were extracted in hexane/acetone and analysed using GC-FID. Polycyclic aromatic hydrocarbons (PAHs) were extracted in hexane/acetone/triethylamine and analysed using GC-MS⁴⁶. Metals were extracted in nitric acid/hydrochloric acid using microwave digestions and analysed using ICP-OES⁴⁷. Nutrients (nitrate, nitrite, phosphate) were measured using an autoanalyser⁴⁸.

Sediment RNA extraction, sequencing and bioinformatics. To mitigate some of the potential issues of relic DNA⁴⁹ within sediments, RNA was used for microbial and meiofaunal community assessment. RNA is thought to give better representation of the “active” fraction of the communities, although this is a subject of debate⁵⁰.

Sediment cores for microbial and meiofaunal analysis were homogenized and weighed (0.25 g) into tubes containing 0.5 g glass beads (100–300 μm , MPBIO) and 1 mL TRI Reagent® (Ambion) before bead beating. Samples were heated at 60°C for 10 min before 600 μL of supernatant was transferred to a new tube containing 100 μL 1-bromo-3-chloro-propane and vortexed. The tubes were centrifuged to separate the organic and aqueous phases before the aqueous phase was transferred to QIA-shredder columns (Qiagen, U.K) containing 0.2 g polyvinylpyrrolidone to remove humic acids and phenolic compounds. The resulting filtrate was precipitated with 2-propanol (equal volume) and sodium acetate (1/10 volume) for 1 hr at -20°C before the RNA pellet was isolated and washed by centrifugation. The RNA pellet was resuspended in 100 μL RNase-free water and further cleaned using the RNeasy kit (Qiagen) according to the manufacturer’s instructions. DNase treatment was performed using RQ1 RNase-Free DNase (Promega) according to the manufacturer’s instructions. Control PCRs confirmed the presence of RNA only. Generation of cDNA was performed using an Omniscript RT kit (Qiagen) in accordance with manufacturer’s instructions.

Bacterial 16S rRNA, archaeal 16S rRNA and eukaryote 18S rRNA gene amplicon sequencing was carried as previously described⁴⁶. Sequencing of amplified bacterial 16S rRNA and eukaryote 18S rRNA genes was performed on an Ion Torrent PGM (Life technologies) and sequencing of amplified archaeal 16S rRNA genes was performed on a MiSeq (Illumina) according to manufacturer’s instructions. Sequences were analysed using the QIIME version 1.8.1 software package and USEARCH version 8. Quality filters were used to remove short (<150 bp) and low quality reads (average Phred score < 25). Chimeras were then identified using UCHIME⁵¹ using the Greengenes database (release 13_5)⁵² for 16S and the SILVA database (version 128)⁵³ for 18S as references. OTUs were clustered using the UPARSE algorithm⁵⁴ defined at 97% similarity for 16S and 18S libraries. For the 18S libraries this similarity threshold was selected as it has been shown to be appropriate for a broad range micro-eukaryote taxonomic groups and specific groups such the Chlorophyta^{55–57}. This cut-off has been shown to be appropriate for both macro/meio-organisms on mock communities^{58,59}. However, it should be noted for other groups 99% (such as some groups of ciliates) may be more appropriate^{60,61}.

Sequences were classified against their respective databases using UCLUST⁶² and OTU tables generated. Eukaryote sequence reads were split into protists using a proposed classification⁴³ and meiofauna based on their taxon identification (e.g. Nematodes, Platyhelminthes, Ostracods and Rotifers). The OTU tables were rarefied to fixed sequence depth for each sample (Bacteria 8000, Archaea 4000, Protists 13000, Meiofauna 1000) and these OTU tables served as input for further statistical tests.

Macrofaunal processing and community analyses. Macrofaunal samples were washed with water over a 500 μm sieve to remove excess sediment. Retained faunal samples were sorted, identified to species level, where possible, and enumerated using a combination of dissecting and compound microscopy. Specimens were dyed to ease identification when necessary using methylene blue ($\text{C}_{16}\text{H}_{18}\text{ClN}_3\text{S}$, CAS: 61-73-4). Identified taxa were labelled and stored in a 70% Industrial Methylated Spirits solution inside glass snap-cap vials. Compilation of the species list was finalised by cross-reference and synonym adjustment consistent with the World Register of Marine Species index (WoRMS; <http://www.marinespecies.org/index.php>). To calculate the BP_c index for each site as described¹⁸, macrofaunal data was averaged across site replicates and macrofaunal wet weights were recorded following blot-drying for <5 seconds. The IQI^3 incorporating the AMBI^2 was estimated for macrofauna data to gauge ecosystem health, following HAPS to GRAB instrument conversion of volumes.

Variable reduction and model fitting. The 49 variables measured at each site were reduced to a limited number of potential predictors of assemblage structure using PCA. Variables were grouped into sets reflecting sediment granulometry, hydrocarbon content, a more generic ‘environmental’ category (representing trace metals and other elements, nutrients and carbon) and a categorization of bioturbation functional types and BP_c . The classification of bioturbators followed the sediment reworking functional types given for taxa¹⁸: surficial modifiers, biodiffusers, upward conveyors, downward conveyors and upward/downward conveyors. Macrofauna were assigned to functional types and relative frequencies of types were used in subsequent PCA.

Benthic assemblages were tested for association against the variable sets summarized in PCA scores using CCA²⁰. CCA plots simultaneously show the dispersion between sites and species based on tables of sites by species, but also express the main axes as linear combinations of predictor variables. Even with variable reduction using PCA, the resultant set of nine predictor variables (eight sets of PCA scores using PC1 and PC2 for each grouped variable set and BP_c values) was sufficient to saturate the CCA for seven sites. The most informative

predictors for ordinations were therefore chosen using forward stepwise selection. At each step, the most informative model was chosen on the basis of a pseudo-F statistic that compares the relative magnitude of the inertia (variance) around the constrained model with predictor(s) to the unconstrained model (no predictors). Significance of the pseudo-F statistic can be assessed by randomly permuting the data with the expectation that an informative predictor will have a lower inertia than the values obtained when there is no pattern in the data. Model building continued by adding predictors one at a time, stopping when additional predictors did not produce a significant reduction in inertia. In some cases, no individual variables provided a significant correlate of observed variation. In these cases, the CCA reverts to an unconstrained analysis, i.e. a correspondence analysis (CA). All ordinations and significance testing were carried out in the vegan package in R⁶³.

The CCA analyses can identify important covariates of variation in taxon identity across sites. The congruence in spatial pattern among taxa can be examined using differences between sites based on taxon turnover (beta diversity) measures, quantified using the presence/absence based Jaccard dissimilarity. Jaccard dissimilarities can be interpreted as the probability that two taxa, drawn from separate sites, will not be shared between the sites⁶⁴. If the different groups of taxa are structured in the same way among sites, dissimilarities between pairs of sites will be related. For example, sites that are relatively similar when measured in terms of archaeal taxa will also share more bacterial taxa than average. This hypothesis can be tested with a Procrustes test, where the minimum difference between ordinations following scaling and rotation is compared to the results of permuting random site composition (the PROTEST statistic)⁶⁵. Ordinations based on Jaccard dissimilarities were made using principal coordinates analysis (PCO) in R. Procrustes statistics were based on all six dimensions from a PCO, as the first two axes only explained approximately 60% of the variation between sites. The Procrustes test is thought to have equal or better power than a Mantel test of the matrix of dissimilarities from one group correlated against dissimilarities from another group⁶⁵. The Mantel test randomises one matrix to establish whether the correlation between a pair of dissimilarity matrices is greater than is likely to have occurred by chance. An advantage of a Mantel correlation is that it is easier to interpret than the PROTEST statistic, where the latter is only equivalent to a correlation when two-dimensional ordinations are superimposed. Mantel statistics (Spearman's rank correlation) are therefore also presented to aid interpretation. The results of Mantel tests and PROTEST were identical in terms of statistical significance.

Data Availability

The data generated and analysed during this study is available as supplementary information to this publication and molecular sequence data can be found at the European Nucleotide Archive (Accession reference code PR-JEB24674). Taxon by site tables (.txt format) can be made available on request.

References

1. Aylagas, E. *et al.* A bacterial community-based index to assess the ecological status of estuarine and coastal environments. *Mar. Pollut. Bull.* **114**, 679–688 (2017).
2. Borja, A., Franco, J. & Pérez, V. A marine biotic index to establish the ecological quality of soft-bottom benthos within European estuarine and coastal environments. *Mar. Pollut. Bull.* **40**, 1100–1114 (2000).
3. Prior, A., Miles, A. C., Sparrow, A. J. & Price, N. Development of a classification scheme for the marine benthic invertebrate component, Water Framework Directive Phase I & II - Transitional and coastal waters. 1–154 (Environment Agency, 2004).
4. Zhong, X., Xu, G. & Xu, H. Use of multiple functional traits of protozoa for bioassessment of marine pollution. *Mar. Pollut. Bull.* **119**, 33–38 (2017).
5. Yazdani Foshtomi, M. *et al.* The link between microbial diversity and nitrogen cycling in marine sediments is modulated by macrofaunal bioturbation. *PLoS One* **10**, e0130116 (2015).
6. Mojtahid, M., Jorissen, F. & Pearson, T. H. Comparison of benthic foraminiferal and macrofaunal responses to organic pollution in the Firth of Clyde (Scotland). *Mar. Pollut. Bull.* **56**, 42–76 (2008).
7. Kennedy, A. D. & Jacoby, C. A. Biological indicators of marine environmental health: meiofauna – a neglected benthic component? *Environ. Monit. Assess.* **54**, 47–68 (1999).
8. Lear, G., Boothroyd, I. K. G., Turner, S. J., Roberts, K. & Lewis, G. D. A comparison of bacteria and benthic invertebrates as indicators of ecological health in streams. *Freshw. Biol.* **54**, 1532–1543 (2009).
9. Kroeker, K. J., Micheli, F., Gambi, M. C. & Martz, T. R. Divergent ecosystem responses within a benthic marine community to ocean acidification. *Proc. Natl. Acad. Sci.* **108**, 14515–14520 (2011).
10. Ellis, R. J. *et al.* Comparison of microbial and meiofaunal community analyses for determining impact of heavy metal contamination. *J. Microbiol. Methods* **45**, 171–185 (2001).
11. Corte, G. N. *et al.* Cross-taxon congruence in benthic communities: Searching for surrogates in marine sediments. *Ecol. Indic.* **78**, 173–182 (2017).
12. Sutherland, T. F., Levings, C. D., Petersen, S. A., Poon, P. & Piercy, B. The use of meiofauna as an indicator of benthic organic enrichment associated with salmonid aquaculture. *Mar. Pollut. Bull.* **54**, 1249–1261 (2007).
13. Pawlowski, J. *et al.* Benthic monitoring of salmon farms in Norway using foraminiferal metabarcoding. *Aquac. Environ. Interact.* **8**, 371–386 (2016).
14. Smale, D. A., Taylor, J. D., Coombs, S. H., Moore, G. & Cunliffe, M. Community responses to seawater warming are conserved across diverse biological groupings and taxonomic resolutions. *Proc. R. Soc. B Biol. Sci.* **284**, 20170534 (2017).
15. Langston, W. J. *et al.* Characterisation of the South West European marine sites: Plymouth Sound and estuaries cSAC, SPA. *Mar. Biol. Assoc.* **9**, 1–156 (2003).
16. Levin, L. A. *et al.* The function of marine critical transition zones and the importance of sediment biodiversity. *Ecosystems* **4**, 430–451 (2001).
17. Covich, A. P. *et al.* The role of biodiversity in the functioning of freshwater and marine benthic ecosystems. *Bioscience* **54**, 767–775 (2004).
18. Queirós, A. M. *et al.* A bioturbation classification of European marine infaunal invertebrates. *Ecol. Evol.* **3**, 3958–3985 (2013).
19. Schwinghamer, P. Characteristic size distributions of integral benthic communities. *Can. J. Fish. Aquat. Sci.* **38**, 1255–1263 (1981).
20. Ter Braak, C. J. F. Canonical correspondence analysis: A new eigenvector technique for multivariate direct gradient analysis. *Ecology* **67**, 1167–1179 (1986).
21. Włodarska-Kowalczyk, M., Górską, B., Deja, K. & Morata, N. Do benthic meiofaunal and macrofaunal communities respond to seasonality in pelagic processes in an Arctic fjord (Kongsfjorden, Spitsbergen)? *Polar Biol.* **39**, 2115–2129 (2016).

22. Gong, J., Dong, J., Liu, X. & Massana, R. Extremely high copy numbers and polymorphisms of the rDNA operon estimated from single cell analysis of oligotrich and peritrich ciliates. *Protist* **164**, 369–379 (2013).
23. Bik, H. M. *et al.* Metagenetic community analysis of microbial eukaryotes illuminates biogeographic patterns in deep-sea and shallow water sediments. *Mol. Ecol.* **21**, 1048–1059 (2012).
24. Massana, R. *et al.* Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. *Environ. Microbiol.* **17**, 4035–4049 (2015).
25. Lennon, J. T., Muscarella, M. E., Placella, S. A. & Lehmkuhl, B. K. How, when, and where relic DNA affects microbial diversity. *MBio* **9**, e00637–18 (2018).
26. Capo, E., Debroas, D., Arnaud, F. & Domaizon, I. Is planktonic diversity well recorded in sedimentary DNA? Toward the reconstruction of past protistan diversity. *Microb. Ecol.* **70**, 865–875 (2015).
27. Forster, D. *et al.* Benthic protists: the under-charted majority. *FEMS Microbiol. Ecol.* **92**, (2016).
28. Leizerowicz, F. *et al.* High-throughput sequencing and morphology perform equally well for benthic monitoring of marine ecosystems. *Sci. Rep.* **5**, 13932 (2015).
29. Lobo, J., Shokralla, S., Costa, M. H., Hajibabaei, M. & Costa, F. O. DNA metabarcoding for high-throughput monitoring of estuarine macrobenthic communities. *Sci. Rep.* **7**, 15618 (2017).
30. Wangenstein, O. S., Palacin, C., Guardiola, M. & Turon, X. DNA metabarcoding of littoral hard-bottom communities: high diversity and database gaps revealed by two molecular markers. *PeerJ* **6**, e4705 (2018).
31. Bik, H. M. *et al.* Sequencing our way towards understanding global eukaryotic biodiversity. *Trends Ecol. Evol.* **27**, 233–243 (2012).
32. Creer, S. *et al.* The ecologist's field guide to sequence-based identification of biodiversity. *Methods Ecol. Evol.* **7**, 1008–1018 (2016).
33. Pawlowski, J. *et al.* The future of biotic indices in the ecogenomic era: Integrating (e) DNA metabarcoding in biological assessment of aquatic ecosystems. *Sci. Total Environ.* **637**, 1295–1310 (2018).
34. Aylagas, E., Borja, Á., Muxika, I. & Rodríguez-Ezpeleta, N. Adapting metabarcoding-based benthic biomonitoring into routine marine ecological status assessment networks. *Ecol. Indic.* **95**, 194–202 (2018).
35. Carrigg, C., Rice, O., Kavanagh, S., Collins, G. & O'Flaherty, V. DNA extraction method affects microbial community profiles from soils and sediment. *Appl. Microbiol. Biotechnol.* **77**, 955–964 (2007).
36. Edgar, R. C. Accuracy of microbial community diversity estimated by closed-and open-reference OTUs. *PeerJ* **5**, e3889 (2017).
37. Cattaneo, A., Asioli, A., Comoli, P. & Manca, M. Organisms' response in a chronically polluted lake supports hypothesized link between stress and size. *Limnol. Oceanogr.* **43**, 1938–1943 (1998).
38. Sulmon, C. *et al.* Abiotic stressors and stress responses: What commonalities appear between species across biological organization levels? *Environ. Pollut.* **202**, 66–77 (2015).
39. Paavola, R. *et al.* Spatial scale affects community concordance among fishes, benthic macroinvertebrates, and bryophytes in streams. *Ecol. Appl.* **16**, 368–379 (2006).
40. Aubé, J. *et al.* The impact of long-term hydrocarbon exposure on the structure, activity, and biogeochemical functioning of microbial mats. *Mar. Pollut. Bull.* **111**, 115–125 (2016).
41. Tait, K. *et al.* Dynamic responses of the benthic bacterial community at the Western English Channel observatory site L4 are driven by deposition of fresh phytodetritus. *Prog. Oceanogr.* **137**, 546–558 (2015).
42. Piot, A., Nozais, C. & Archambault, P. Meiofauna affect the macrobenthic biodiversity-ecosystem functioning relationship. *Oikos* **123**, 203–213 (2014).
43. Caron, D. A., Countway, P. D., Jones, A. C., Kim, D. Y. & Schnetzer, A. Marine protistan diversity. *Ann. Rev. Mar. Sci.* **4**, 467–493 (2012).
44. Needham, H. R., Pilditch, C. A., Lohrer, A. M. & Thrush, S. F. Context-specific bioturbation mediates changes to ecosystem functioning. *Ecosystems* **14**, 1096–1109 (2011).
45. Blott, S. J. & Pye, K. Gradistat: A grain size distribution and statistics package for the analysis of unconsolidated sediments. *Earth Surf. Process. Landforms* **26**, 1237–1248 (2001).
46. Taylor, J. D. & Cunliffe, M. Polychaete burrows harbour distinct microbial communities in oil-contaminated coastal sediments. *Environ. Microbiol. Rep.* **7**, 606–613 (2015).
47. Sun, W. P. *et al.* Sources and geographic heterogeneity of trace metals in the: Sediments of Prydz Bay, East Antarctica. *Polar Res.* **32**, 20049 (2013).
48. Widdicombe, S. & Needham, H. R. Impact of CO₂-induced seawater acidification on the burrowing activity of *Nereis virens* and sediment nutrient flux. *Mar. Ecol. Prog. Ser.* **341**, 111–122 (2007).
49. Carini, P. *et al.* Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. *Nat. Microbiol.* **2**, 16242 (2017).
50. Blazewicz, S. J., Barnard, R. L., Daly, R. A. & Firestone, M. K. Evaluating rRNA as an indicator of microbial activity in environmental communities: Limitations and uses. *ISME J.* **7**, 2061–2068 (2013).
51. Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. & Knight, R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**, 2194–2200 (2011).
52. DeSantis, T. Z. *et al.* Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72**, 5069–5072 (2006).
53. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**, 590–596 (2013).
54. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* **10**, 996 (2013).
55. Amaral-Zettler, L. A., McCliment, E. A., Ducklow, H. W. & Huse, S. M. A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PLoS One* **4**, e6372 (2009).
56. Taylor, J. D. & Cunliffe, M. High-throughput sequencing reveals neustonic and planktonic microbial eukaryote diversity in coastal waters. *J. Phycol.* **50**, 960–965 (2009).
57. Tragin, M., Zingone, A. & Vault, D. Comparison of coastal phytoplankton composition estimated from the V4 and V9 regions of the 18S rRNA gene with a focus on photosynthetic groups and especially Chlorophyta. *Environ. Microbiol.* **20**, 506–520 (2018).
58. Brannock, P. M. & Halanych, K. M. Meiofaunal community analysis by high-throughput sequencing: Comparison of extraction, quality filtering, and clustering methods. *Mar. Genomics* **23**, 67–75 (2015).
59. Brown, E. A., Chain, F. J. J., Crease, T. J., Macisaac, H. J. & Cristescu, M. E. Divergence thresholds and divergent biodiversity estimates: Can metabarcoding reliably describe zooplankton communities? *Ecol. Evol.* **5**, 2234–2251 (2015).
60. Bachy, C., Dolan, J. R., López-García, P., Deschamps, P. & Moreira, D. Accuracy of protist diversity assessments: Morphology compared with cloning and direct pyrosequencing of 18S rRNA genes and ITS regions using the conspicuous tintinnid ciliates as a case study. *ISME J.* **7**, 244–255 (2013).
61. Santoferrara, L. F., Grattepanche, J. D., Katz, L. A. & Mcmanus, G. B. Pyrosequencing for assessing diversity of eukaryotic microbes: Analysis of data on marine planktonic ciliates and comparison with traditional methods. *Environ. Microbiol.* **16**, 2752–2763 (2014).
62. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
63. Oksanen, J. *et al.* Package 'vegan'. *R Packag. ver. 2.0-8* 254 (2016).
64. Anderson, M. J., Ellingsen, K. E. & McArdle, B. H. Multivariate dispersion as a measure of beta diversity. *Ecol. Lett.* **9**, 683–693 (2006).
65. Peres-Neto, P. R. & Jackson, D. A. How well do multivariate data sets match? The advantages of a procrustean superimposition approach over the Mantel test. *Oecologia* **129**, 169–178 (2001).

Acknowledgements

A special thank you to Dr. Robert Kennedy of the National University of Ireland, Galway, Dr. Jack O'Carroll and Dr. Adrian Patterson for their assistance during macrofaunal and sediment analyses. This research was financially supported by the European Union 7th Framework Programme for Research and Technological Development under GA No.312139, as part of the Integrated Biotechnological Solutions for Combating Marine Oil Spills project Kill Spill.

Author Contributions

S. Cronin-O'Reilly: Macrofaunal data acquisition, data analysis and manuscript preparation. J.D. Taylor: Field survey design and implementation, microbial community analysis, bioinformatics analysis, environmental data acquisition and manuscript preparation. I. Jermyn: Field survey design and implementation, sediment data acquisition and manuscript preparation. A.L. Allcock: Qualitative assessment and supervision of field survey, macrofaunal community and sediment processing. Manuscript preparation. M. Cunliffe: Microbial community data analysis and manuscript preparation. M.P. Johnson: Data analysis and interpretation and manuscript preparation.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-33799-9>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018