Vol. 735: 1–26, 2024
https://doi.org/10.3354/meps14576

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

Published May 2

FEATURE ARTICLE

Macrofaunal diversity patterns in coastal marine sediments: re-examining common metrics and methods

Mary E. Clinton[1,*], Paul V. R. Snelgrove[1,2], Amanda E. Bates[1,3]

[1]Department of Ocean Sciences, Memorial University of Newfoundland, St. John's, Newfoundland and Labrador A1C 5S7, Canada
[2]Department of Biology, Memorial University of Newfoundland, St. John's, Newfoundland and Labrador A1B 3X9, Canada
[3]Department of Biology, University of Victoria, Victoria, British Columbia V8P 5C2, Canada

ABSTRACT: Complex biodiversity patterns arise in marine systems due to overlapping ecological processes, including organism interactions, resource distribution, and environmental conditions. Despite the importance of documenting these patterns, describing diversity in natural ecosystems remains challenging. Here, we investigate 3 nearshore sub-Arctic sites to describe benthic macroinfaunal taxa and biological traits, with the ultimate aim of determining whether common diversity metrics and typical sampling efforts adequately capture community composition in these systems. First, we assess how diversity relates to sediment depth and examine relationships among commonly used taxonomic and functional diversity indices. Second, using a power analysis, we explore how sampling effort influences the interpretation of diversity patterns in coastal systems. We report significant variation in community composition among sites, even across small spatial scales of km, and find that taxonomically diverse communities do not necessarily correspond to high functional diversity. We further find that although environmental factors such as sediment depth consistently affect macroinfaunal diversity, the direction and magnitude of these relationships are site-dependent. Finally, we demonstrate that typical sampling effort for coastal benthic studies (for example, <5 push cores of ~7 cm diameter) may not capture macroinfaunal community composition adequately, potentially obscuring hotspots in common diversity metrics such as taxonomic or functional richness. However, indices such as Simpson's diversity may be well-suited to resource-limited studies with restricted sampling capacity. Our results highlight the importance of adopting multi-pronged approaches to biodiversity assessment and determining optimal sample sizes for marine benthic systems, particularly in the context of biodiversity monitoring for conservation purposes.

A subset of the diverse macroinvertebrates inhabiting nearshore sub-Arctic sediments, in which sampling effort can drastically alter inferences on diversity patterns.

Photo/Graphic: Mary E. Clinton

KEY WORDS: Biodiversity · Benthic infauna · Seafloor · Functional traits · Power analysis

1. INTRODUCTION

Biological diversity influences the health, functioning, and stability of natural ecosystems and, by extension, human wellbeing (Cardinale et al. 2012, Naeem et al. 2016). As a result, biodiversity measurement defines a central pillar of conservation frameworks

*Corresponding author: meclinton@mun.ca

Publisher: Inter-Research · www.int-res.com

and monitoring programs globally (Wabnitz et al. 2010, Convention on Biological Diversity 2022). However, the complexity of natural ecosystems leaves accurate quantification of diversity as one of the primary ongoing challenges of ecological research (Magurran 1988, Purvis & Hector 2000). Indeed, many overlapping processes influence community composition, including organism interactions (e.g. competition, predation, reproduction), disturbance, and resource distribution (Legendre & Fortin 1989, Collinge 2001, Rietkerk & van de Koppel 2008). In marine ecosystems, environmental conditions such as depth, salinity, and organic matter availability interact and influence diversity as well (Ellingsen 2002, Baldrighi et al. 2014, Campanyà-Llovet et al. 2017). The interplay of these processes produces complex biodiversity patterns that often vary widely, even over relatively small spatial scales (Seed 1996, Silberberger et al. 2018). As such, accurate description of biodiversity patterns or investigation of diversity–environment relationships requires careful consideration of appropriate metrics and methodologies.

Historically, researchers have quantified biodiversity using metrics of taxonomic composition, such as species richness and evenness, Shannon-Wiener index, and Simpson's diversity, which describe the number of taxa present in a community and their relative abundances (Shannon 1948, Simpson 1949, Pielou 1975). However, ecologists increasingly recognize the importance of functional diversity, which considers the distribution of trait values within a community (Tilman 2001, Petchey & Gaston 2002, Miatta et al. 2021). Biological traits include morphological, physiological, and behavioural characteristics of organisms — such as body size, mobility, and feeding mode — that jointly determine the ecological roles of organisms and how they interact with their environment and with one another (Violle et al. 2007, Lefcheck et al. 2015). Functional diversity indices describe the range, prevalence, and distribution of these traits within an ecosystem (Díaz & Cabido 2001) and are widely considered better predictors of ecosystem functioning than their taxonomic counterparts alone (Díaz & Cabido 2001, Tilman 2001, Bremner et al. 2003, Hooper et al. 2005, Mouillot et al. 2011, Lefcheck & Duffy 2015). However, quantifying functional diversity requires empirical data for multiple traits across large numbers of taxa (Lefcheck et al. 2015), which presents challenges in many data-deficient marine environments. Polar marine ecosystems are particularly challenging in this regard, given the logistical difficulties of sample collection, lack of trait knowledge for many high-latitude species, and poorly understood trait–function relationships (Degen et al. 2018). Despite these challenges, the recent development of trait databases (e.g. Faulwetter et al. 2014, Degen & Faulwetter 2019) and ongoing efforts to share best practices, methodology, and terminology amongst researchers continues to promote the expansion of trait-based studies in high-latitude environments.

Coastal macroinfaunal communities provide an ideal study system for examining questions related to changes in marine biodiversity. Benthic macrofauna are highly abundant, taxonomically diverse, and encompass a wide range of feeding modes and life-history strategies (Pinto et al. 2009, Patrício et al. 2012). Additionally, nearshore environments enable relatively simple and inexpensive benthic sample collection. Unlike studies of deep-sea systems, where logistical constraints such as ship availability or equipment costs limit sampling, coastal assessments of marine benthos provide a valuable opportunity to examine the potential effects of sampling effort on the interpretation of diversity studies. The availability of taxonomic keys and functional trait databases (Faulwetter et al. 2014, Fofonoff et al. 2018, Degen & Faulwetter 2019) for many coastal taxa adds further advantages.

Here, we examine sedimentary biodiversity patterns in 3 nearshore sites on the sub-Arctic island of Newfoundland, Canada, to achieve 4 key research outcomes. (1) We describe each biological community by producing a comprehensive taxonomic list of the macroinfaunal invertebrates present at each site, including descriptions of key biological traits. (2) We then calculate a suite of taxonomic and functional diversity indices and examine the relationship between sediment depth and macroinfaunal diversity at each of these 3 sites. (3) Next, we use data from all sites to test for consistent relationships between commonly used taxonomic and functional diversity indices. (4) Finally, we take advantage of the unique opportunity provided by the comparatively large number of replicate samples — relative to most benthic studies — to examine how replication influences interpretation of diversity patterns in seafloor ecosystems.

2. MATERIALS AND METHODS

2.1. Study sites and field sampling

We collected sediment push cores in 3 distinct coastal environments on the island of Newfoundland, Canada (Fig. 1). St. Paul's Bay (SP; 49.864° N, −57.815° W), a sandy tidal flat located just off the Gulf of St. Lawrence on the west coast of Newfound-

land, experiences seasonal sea ice cover. Previous biodiversity assessments suggest a macrofaunal transition in the region between the Labrador and Acadian biogeographic provinces (Quijón & Snelgrove 2005). The second site, Neddie's Harbour (NH), is also located on the west coast of Newfoundland, south of SP (49.524° N, −57.884° W). This site is located in the East Arm of Bonne Bay, a sub-Arctic fjord; a shallow sill (~15 m deep) separates the East Arm from the South Arm and the Gulf of St. Lawrence. Finally, Newman Sound (NS; 48.557° N, −53.965° W) is a glacial fjord in Terra Nova National Park on the eastern coast of Newfoundland, comprising inner and outer sounds separated by a sill located 7 km from the head of the fjord; our study site is located within the inner sound. In contrast to the west coast sites, the macrofaunal community in NS does not experience sea ice in the winter months, presumably contributing to a more temperate fauna.

At each study site, we waded into the shallow subtidal and collected sediment push cores (outer diameter 7.3 cm, inner diameter 6.7 cm) by hand between September and October 2020. We collected 33 cores at SP and NS; however, equipment constraints limited us to 24 cores at NH. Cores retained at least 10 cm of sediment and were collected ~30 cm apart from one another over a relatively small area (<100 m^2) at a distance of ~10 m from the high tide line. We maintained cores immersed in seawater and at ambient temperatures until processing to minimize vertical movement of fauna but cannot rule out the possibility of some activity. To minimize disruption of sediment layers, we sealed the cores with large rubber stoppers and transported them to the laboratory in shock-absorbing, insulated carrying totes.

2.2. Environmental properties

At each study site, we collected one push core to measure sediment properties (in addition to cores for

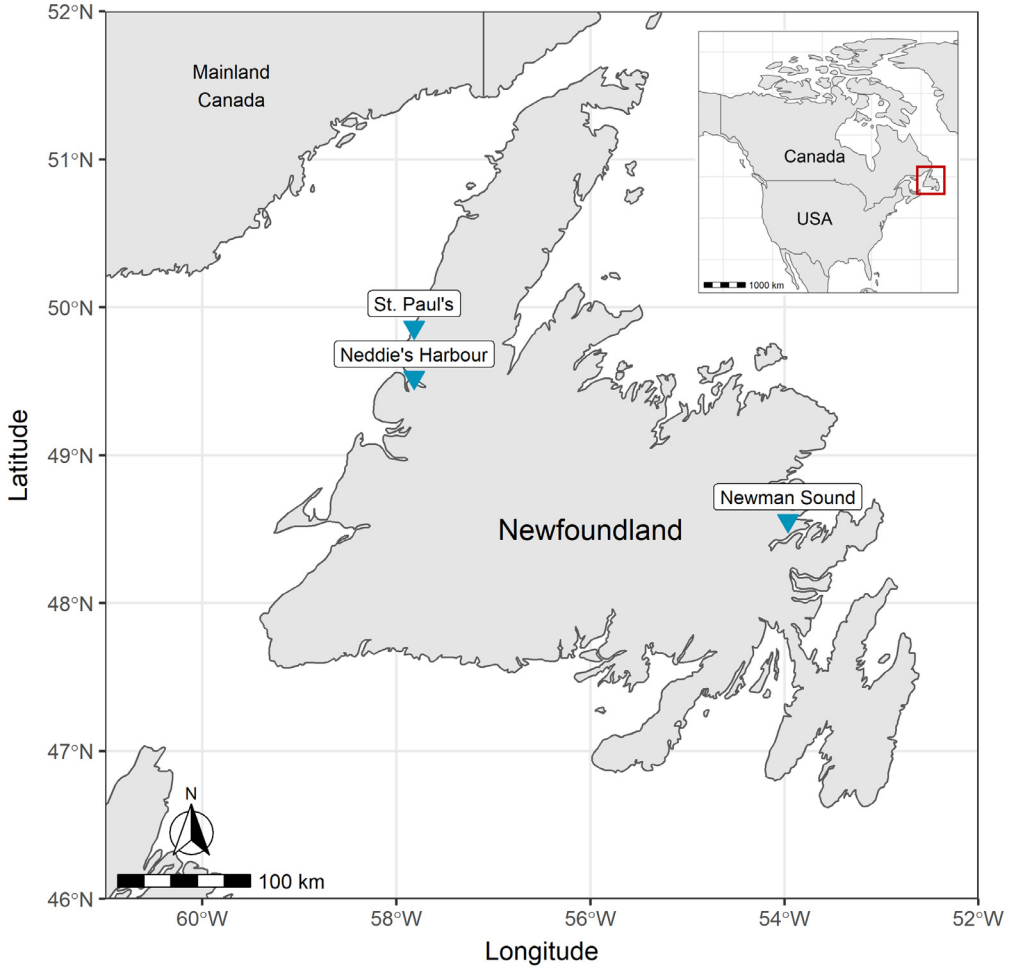

Fig. 1. Sites in Newfoundland, Canada. Inset shows North America, with the red box outlining the island of Newfoundland. Sediment push cores were collected by hand at St. Paul's Bay, Neddie's Harbour, and Newman Sound

species characterization); equipment constraints precluded the possibility of additional replicates. We extruded and sectioned each of these cores into 3 depth layers (0–2, 2–5, 5–10 cm) using an inert plastic spatula. For each depth layer, we assessed granulometric properties by transferring ~5 ml of homogenized sediment into a 50 ml tube and digesting any organic particles in the sample with 35% hydrogen peroxide. We then freeze-dried each sample, sieved them to remove any coarse material (i.e. gravel fraction >2 mm), and analyzed the remaining sediment using a Beckman Coulter LS13-320 laser diffraction analyzer. We logged and processed the laser files using Femto Particle Sizing Software (PSS, version 5.6). For each sample, we determined the relative percentage (%) of gravel (>2 mm), sand (63 µm–2 mm), silt (10–63 µm), and clay (<10 µm), as well as the mean grain size (MGS) of the sortable silt fraction (µm).

Subsamples were collected from the sediment property core to assess organic carbon (%C) as a measure of food quantity in combination with nitrogen (%N) and carbon-to-nitrogen ratio (C:N), as indicators of longer term food quality at each site. We considered higher %N and lower C:N indicative of fresher and higher quality organic matter (Godbold & Solan 2009, Le Guitton et al. 2015, Campanyà-Llovet et al. 2017). We froze these samples in Whirl-Pak bags at −20°C until analysis. We then dried sediment at 80°C for 24 h, homogenized with a ball grinder, and transferred samples to a desiccator with an open beaker of HCl to remove inorganic carbon (Danovaro 2010). We allowed sediment to dry for another 24 h before subsampling material into tin capsules for analysis using a Vario Isotope Cube (Elementar) elemental analyzer.

We also collected 3 mini-cores (50 ml syringes with syringe barrel tips removed, inner diameter 2.7 cm) from each study site and determined concentrations of chlorophyll *a* (chl *a*) and phaeopigments (Phaeo) in the surface layer of each sample following Danovaro (2010). We placed 1–2 g of sediment in a pre-weighed 15 ml tube containing ~8 ml of 90% acetone, and vortexed and sonicated each sample before storing samples overnight in the dark. The following day, we centrifuged each sample, transferred 3 ml of supernatant to a cuvette, and used a spectrophotometer to measure the absorbance of the supernatants at 2 wavelengths (750 and 665 nm). We added 200 µl of 0.1 N HCl to the supernatant before measuring absorbance again at the same wavelengths. We then dried sediments under a fume hood for 10 d before reweighing each sample, allowing us to standardize pigment concentrations per gram of sediment. Chl *a*:Phaeo ratios provide an indication of organic matter quality over the short term, with higher ratios suggesting fresher material (Le Guitton et al. 2015).

2.3. Faunal push core processing

To assess taxonomic composition of the macroinfaunal community, we extruded and sectioned each core into 3 depth layers (0–2, 2–5, 5–10 cm) using an inert plastic spatula. We then sieved the sediment through a 500 µm mesh and fixed each depth layer in a 4% buffered seawater–formaldehyde solution, minimizing handling to protect fragile macrofauna from damage. In the laboratory, we transferred sediment samples to a 70% ethanol solution for longer term storage, until we could begin microscopic analyses. To facilitate infauna identification, we stained each sample with a few drops of Rose Bengal (0.5 g l^{-1}) before sorting the sediment under a dissecting microscope and visually identifying macrofauna to the lowest possible taxonomic level. In some cases, taxonomic resolution differed among sites for a given taxonomic unit. To avoid biases resulting from these differing levels of taxonomic uncertainty, we binned some taxa before calculating diversity indices for each site (see the Appendix).

2.3.1. Assessment of taxonomic diversity

For each site, we counted the total number of individuals of each taxon and calculated the following indices of taxonomic diversity for each core: taxonomic richness (*S*; defined as the number of taxa present in each sediment core), Shannon-Wiener (*H′*; natural logarithm), Simpson's index (*d*), and Pielou's evenness (*J′*). These indices were calculated in R using the 'vegan' package (Oksanen et al. 2020, R Core Team 2021). Finally, to assess the adequacy of our sampling efforts, we constructed species accumulation curves to illustrate the total number of macrofaunal taxa observed as a function of the number of push cores collected at each site.

We also determined the total number of rare taxa observed at each site. We define rare taxa as those represented by no more than 5 individuals in total across all samples collected from a given site, and which were observed in <10% of cores at that site; for SP and NS, this rarity threshold equates to ~0.15 individuals per sediment core. In NH, where we only collected 24 cores, we adjusted the threshold to account for differences in sampling effort. Rare taxa in NH

therefore include those which were observed in <10% of cores and which were represented by no more than 4 individuals in total. We compared both threshold approaches for NH and found that the 2 methods yielded identical lists of rare taxa.

2.3.2. Assessment of functional diversity

To evaluate functional composition of the macrofaunal communities, we selected 5 functional traits based on their relationship to ecosystem functioning and availability of data for all taxa. The selected traits reflect both morphology (body size) and behaviour (diet, feeding guild, mobility, sediment reworking) (see Table 1). Published sources (Fauchald & Jumars 1979, Macdonald et al. 2010, Queirós et al. 2013, Fofonoff et al. 2018, Degen & Faulwetter 2019, Polytraits Team 2022, WoRMS Editorial Board 2022) provided trait information for each taxon collected. When trait information was unavailable for a specific taxon, we obtained information from the next taxonomic rank (typically genus level) with known trait data.

We determined body size for each taxon based on the largest attainable size reported in the literature. We then log transformed body size and scaled the resulting values between 0 and 1 to ensure this variable did not disproportionately affect Gower's distances. Diet and feeding guild are nominal trait variables with 3 and 6 modalities, respectively. We used a fuzzy coding approach for these traits to allow for intraspecific variation in trait expression and scored each taxon from 0 to 1 depending on the extent to which a taxon expressed a given trait. For example, the spionid worm *Pygospio elegans* obtains food via suspension feeding or surface deposit feeding, depending on flow conditions; therefore, we assigned this taxon a feeding guild value of 0.5 for each of these trait modalities. Given that mobility is an ordinal variable, we assigned each taxon a value of 0 (none), 0.3 (low), 0.6 (moderate), or 1 (high), with 0 representing a completely sessile taxon and 1 indicating high mobility (Swift 1993). Finally, for the nominal variable sediment reworking, we assigned each taxon to one of 4 categories: epipelagic, surficial modifier, upward or downward conveyor, or biodiffusor (Table 1).

Following assignment of trait modalities to all taxa, we calculated a Gower's distance matrix, providing the overall functional distance between each pair of taxa (Dray & Dufour 2007). We then applied a principal coordinate analysis on this distance matrix to obtain coordinates for each taxon in a multidimensional functional space, following (Villéger et al. 2008). Based on this functional space, we determined the following functional diversity indices: functional richness (FRic), functional evenness (FEve), and functional divergence (FDiv) (Mason et al. 2005). FRic

Table 1. Functional traits, modalities, and definitions. As a continuous variable, body size has no modalities (N/A: not applicable). Definitions for mobility and sediment reworking follow (Queirós et al. 2013)

Trait	Modalities	Definitions of trait modalities
Body size	N/A	Log-transformed maximum body size, scaled between 0 and 1
Diet	Herbivore	Feeds on primary producers
	Carnivore	Feeds on other animals
	Detritivore	Feeds on dead organic matter
Feeding guild	Predator	Actively hunt live prey
	Scavenger	Consumes mostly decaying biomass
	Grazer	Grazes plant material
	Suspension or filter feeder	Filters food particles from the water column
	Sub-surface deposit feeder	Consume particles at depth
	Surface deposit feeder	Collect and consume surface sediment particles
Mobility	None	Sedentary or only moving within a fixed tube structure
	Low	Limited movement (i.e. withdraws into sediment when disturbed)
	Moderate	Slow, free movement through sediment matrix via non-permanent burrow formation
	High	Free, 3-dimensional movement via permanent, excavated burrow system
Sediment reworking	Epipelagic	Epifauna that bioturbate at the sediment surface with no significant contribution to sediment reworking
	Surficial modifier	Activities cause movement of particles at the sediment surface
	Upward or downward conveyor	Feed at depth and defecate at surface, or feed at surface and defecate at depth
	Biodiffusor	Causes constant and random local sediment biomixing over short distances

represents the total functional space occupied by the community (Villéger et al. 2008), quantified as the volume of the convex hull (Cornwell et al. 2006). FEve, a unitless metric constrained between 0 and 1, describes the extent to which faunal abundances are evenly distributed within the trait space, where a value of 1 represents a perfectly even community (Villéger et al. 2008). FDiv quantifies heterogeneity of traits by representing the probability that 2 taxa picked at random from the community exhibit the same trait value (Lavorel et al. 2008). This metric is weighted by taxon abundance; a low FDiv indicates that abundant taxa exhibit functional traits located close to the center of the trait space and a high FDiv indicates extreme trait values for the most abundant taxa (Villéger et al. 2008).

In addition to traditional metrics of functional diversity, we also determined community-level weighted means (CWMs) for each trait. For categorical traits, such as diet or sediment reworking, CWM represents the probability that an individual organism drawn at random from the community will exhibit a given trait modality (Ricotta & Moretti 2011). For example, a CWM of 0.8 for the trait modality 'herbivore' indicates an 80% probability that a random individual drawn from that community would be herbivorous. For continuous traits, such as body size, CWM simply represents the average value for macrofauna in that community (Lavorel et al. 2008). We calculated CWMs for each core and each depth layer within each core and determined site-level values by averaging all cores within each site. Calculations of functional diversity indices and CWMs used the 'FD' package (Laliberté & Legendre 2010, Laliberté et al. 2014) and the functions 'multidimFD' and 'qual_funct_space' (see Villéger et al. 2008, Mouillot et al. 2013 for details).

2.4. Effect of sediment depth on diversity

To determine the effect of sediment depth on macrofaunal diversity, we first assessed the vertical distribution of macrofauna within the sediment by calculating the proportion of fauna present in each depth layer at each site. Next, we calculated taxonomic and functional diversity indices for each depth layer within each core. Preliminary analysis revealed similar diversity patterns in the 2–5 and 5–10 cm depth layers; therefore, we pooled these layers to calculate diversity indices. We also pooled CWMs for the 2–5 and 5–10 cm depth layers in each core, following the methods described above.

2.5. Relationships between taxonomic and functional diversity indices

For each pair of diversity indices, we created scatterplots to identify visually consistent relationships between metrics across all sites. We also calculated Pearson correlation coefficients (r) for each pair.

2.6. Statistical analyses

All statistical analyses were conducted using R statistical software, version 4.1.1 (R Core Team 2021). To test for differences in taxonomic composition between collection dates, we conducted analyses of similarity (ANOSIMs) for each site. Given the negligible effect of collection date on macroinfaunal community composition (see Table S2), we pooled cores from all collection dates within each study site for all subsequent analyses. We used non-metric multidimensional scaling (nMDS) ordination to visualize overall differences in taxonomic composition and diversity indices among sites and conducted permutational analyses of variance (PERMANOVAs) using Bray-Curtis dissimilarity matrices from non-transformed abundance data to test for statistical differences in the centroids of each group (site). We then conducted a similarity of percentages (SIMPER) analysis to determine which taxa contributed most to dissimilarity among sites. PERMANOVA, ANOSIM, SIMPER, and nMDS analyses used the 'vegan' package (Oksanen et al. 2020). To compare diversity indices and CWMs between depth layers within a given site, we performed Wilcoxon signed-rank tests using the 'stats' package (R Core Team 2021).

We modeled each taxonomic and functional diversity index as a function of site, depth layer, and their interaction using linear mixed-effects models via the 'lme4' package (Bates et al. 2015), and calculated parameter-specific p-values using a Kenward-Roger approximation implemented in the 'pbkrtest' package (Halekoh & Højsgaard 2014). For those indices naturally bounded between 0 and 1, we used beta mixed-effects models from the 'glmmTMB' package (Kieschnick & McCullough 2003, Ferrari & Cribari-Neto 2004, Brooks et al. 2017). To account for variability in each index at the sediment core level, we also included core ID as a random intercept. For interaction terms with $p > 0.05$, we re-ran the model without an interaction term. We used the 'DHARMa' package for model diagnostics and to test model fit (Hartig 2022). Although beta regression was the best fit for the *d* index data, DHARMa indicated some borderline

violations of model assumptions. Therefore, we performed non-parametric Kruskal-Wallis tests for *d* as a function of site ($\chi^2 = 56.0$, $p = 6.9 \times 10^{-13}$) and depth layer ($\chi^2 = 0.08$, $p = 0.77$), respectively, which confirmed robust biological inferences drawn from the beta regression model.

2.7. Power analysis

In a statistical context, power refers to the probability that a test of significance — in this case, generalized linear mixed effects models for each diversity index — will detect an effect (i.e. a deviation from the null hypothesis) that is present in the data, should one exist (Steidl & Thomas 2001). To determine whether power would have decreased, and consequently whether our results and inferences may have differed had we collected fewer cores, we took 10 000 random samples (without replacement) of 3, 5, 10, 15, and 20 cores from each site. For each of these sample sizes, we then calculated the mean values and standard deviations for taxonomic and functional diversity indices for each site and for each depth layer within site to explore how diversity indices changed when sampling different numbers of cores.

We then fit 1000 generalized linear mixed effects models for each diversity index, based on 1000 random samples for each sample size (3, 5, 10, 15, and 20 cores). From these model results, we calculated power by determining the proportion of iterations in which each predictor variable (depth layer, site, and their interaction) significantly affected the diversity index ($p < 0.05$). We used 0.80 (or 80%) to indicate a power threshold sufficiently robust for a given sample size; this value corresponds to the minimum conventional power threshold typically used for scientific experiments (Lipsey 1990). All data and code required to reproduce these analyses are available on Dryad (https://doi.org/10.5061/dryad.76hdr7t3k).

3. RESULTS

3.1. Sedimentary characteristics

Based on visual inspection, grain size appeared similar among all sediment cores within each site. All 3 study sites were predominantly sandy, with small amounts of gravel, silt, and clay (Table 2). However, the sediment in SP contained less gravel and a greater percentage of sand than the other 2 sites. SP also contained the smallest MGS of the sortable silt fraction, followed by NH (Table 2). Despite the similar grain size observed visually across all cores, and the modest spatial scales spanned within our sampling locations (<100 m^2), we acknowledge that the single replicates we analyzed for grain size may have missed some within-site variation in sediment composition.

With respect to organic matter concentrations, organic carbon (%C) was ~4 times higher in NH than in NS and SP, resulting in a relatively high C:N ratio at this site (Table 2). Total nitrogen (%N) was similar among sites. Finally, the chl *a*:Phaeo ratio was highest at SP (Table 2). Although these differences are potentially biologically significant, low sample size precluded statistical comparison of nutrient and grain size measurements among sites.

Table 2. Sediment properties (granulometric, organic matter, and photopigment concentrations) for each depth layer at each site (NH: Neddie's Harbour; NS: Newman Sound; SP: St. Paul's). Mean grain size (MGS) indicates size of the sortable silt fraction. Duplicate measures of organic carbon (%C; mg per 100 mg dry weight [DW]) and nitrogen (%N; mg per 100 mg DW) were taken based on sub-samples from a single replicate for each depth layer and used to calculate a carbon to nitrogen ratio (C:N). Standard deviations for %C and %N capture instrumentation variation rather than within-site variation. Chlorophyll *a* (chl *a*; μg g^{-1} DW) and phaeopigment (Phaeo; μg g^{-1} DW) were measured for surface sediment only; standard deviations are based on 3 mini-cores collected at each site

Site	Depth layer (cm)	Gravel (%)	Sand (%)	Silt (%)	Clay (%)	MGS (μm)	%C	%N	C:N	Chl *a*	Phaeo	Chl *a*: Phaeo
NH	0–2	5.4	86.7	5.5	2.4	34.2	0.98 ± 0.031	0.034 ± 0.00028	28.5	2.54 ± 0.29	9.08 ± 0.30	0.279
	2–5	2.0	91.0	4.9	2.1	34.6	0.82 ± 0.013	0.021 ± 0.0016	38.6			
	5–10	4.6	88.6	4.7	2.1	30.8	0.77 ± 0.0085	0.023 ± 0.00078	33.3			
NS	0–2	2.4	91.3	4.9	1.4	36.6	0.29 ± 0.026	0.033 ± 0.0023	8.78	2.71 ± 0.28	9.68 ± 0.69	0.281
	2–5	0.4	93.5	4.8	1.3	41.9	0.14 ± 0.0078	0.016 ± 0.00057	8.66			
	5–10	7.0	88.5	3.5	1.0	37.6	0.20 ± 0.0085	0.020 ± 0.00085	9.80			
SP	0–2	0.4	96.1	2.5	1.0	32.9	0.21 ± 0.0021	0.026 ± 0.00021	8.20	4.23 ± 0.92	9.43 ± 0.52	0.447
	2–5	1.5	94.7	2.7	1.1	33.0	0.21 ± 0.0035	0.026 ± 0.00042	8.28			
	5–10	4.3	92.3	2.4	1.0	32.8	0.18 ± 0.00071	0.021 ± 0.000071	8.89			

3.2. Community overview

In total, we identified 16 357 individuals (Fig. 2, Table S1 in the Supplement at www.int-res.com/articles/suppl/m735p001_supp.pdf). Removal of low-resolution taxonomic groupings and binning of some taxa to standardize taxonomic uncertainty (see the Appendix) reduced that number to 10 048 individuals representing 53 taxa, 12 of which were present at all 3 study sites (Table S1). Together, these taxa represent 4 phyla, 17 orders, and 30 families of macroinfauna. We performed all relevant analyses with both binned (i.e. binned with low-resolution taxonomic groupings removed) and unbinned (i.e. raw) taxonomic data. Given that results differed only marginally and did not alter ecological conclusions, we present only binned data here (see the Appendix).

Bivalves dominated SP, with 2656 individuals of the amethyst gem clam *Gemma gemma*, the most abundant taxon, representing ~52% of the total macrofauna at the site. By contrast, polychaetes dominated NS; the most abundant taxon at this site, *Pygospio elegans* (n = 1338), made up ~ 51% of the macrofaunal community. Finally, several taxonomic groups dominated NH, led by *P. elegans* (n = 463), *Monocorophium* sp. (n = 438), and *Mya* sp. (n = 336), representing ~20, 19, and 15% of total macrofauna, respectively. NS contained the greatest number of taxa, with the highest overall macrofaunal abundance at SP.

Taxonomic composition differed significantly among sites (PERMANOVA, pseudo-R = 64.4, $p < 0.001$; Fig. 3A). Twelve taxa explained a large percentage (~90%) of this dissimilarity (Table 3, SIMPER analysis). Collection date had a negligible effect on taxonomic composition (Table S2), so we pooled all samples within each site across sampling dates.

3.2.1. Taxonomic diversity

When considered together, taxonomic diversity (i.e. S, H', d, and J') differed significantly among sites (PERMANOVA, pseudo-R = 7.41, $p < 0.001$; Fig. 3B, Table S3). We observed the greatest number of taxa at NS, followed by SP and then NH (Fig. 4). However, the diversity of individual cores yielded the reverse pattern, with the highest core-level S at NH, followed by SP, then NS (Fig. 5A). This pattern may reflect the larger number of rare taxa at NS relative to SP and NH (Fig. S1). J' was similar between NH and NS, but lowest in SP (Fig. 5B). Finally, H' and d were the highest and least variable in NH (Fig. 5).

3.2.2. Functional diversity

When considered together, functional diversity (a combination of FRic, FEve, and FDiv) also differed significantly among sites, and this difference appeared more pronounced than that of taxonomic diversity (PERMANOVA, pseudo-R^2 = 27.3, $p < 0.001$; Fig. 3C). Each functional diversity index was highest at a different site, with highest FRic in NH, highest FEve in NS, and highest FDiv in SP (Fig. 5).

To provide insight into which traits might drive these differences in functional diversity indices, we also investigated CWMs at each site. Although organisms were slightly smaller on average in NS compared to the west coast sites, body size was similar among sites (Fig. 6, Table S4). Herbivory dominated the diets of most individuals at all sites (75, 61, and 62% for NH, NS, and SP, respectively; Fig. 6). However, the feeding guild CWMs differed among sites. The most common feeding guild in NS, surface deposit feeders, represented 35% of all macrofauna; this feeding guild was less common in NH (17%) and only minimally represented in SP (~5%). The most abundant surface deposit feeders in NS were *P. elegans*, Tellinidae indet., and Rissooidea indet. By contrast, suspension feeders (e.g. *G. gemma*) dominated SP and grazers (e.g. Naididae indet. and *Littorina* sp.) dominated NH, accounting for 53 and 49% of the total macrofauna at each site, respectively.

Differences in CWMs between the east and west coast sites were even more pronounced with respect to sediment reworking. Surficial modifiers strongly dominated NH and SP, representing 75 and 70% of macrofauna at each site, respectively. Conversely, NS contained far fewer surficial modifiers (28%), instead dominated by upward and downward conveyors such as *P. elegans* and *Polydora cornuta* (52%). Finally, average mobility was notably lower in NS than in the other 2 sites; the prevalence of highly mobile taxa such as *Monocorophium* sp. contributed to the greatest mobility in NH (Fig. 6).

3.3. Relationship between sediment depth and macrofaunal diversity

At all 3 sites, the majority of macrofauna (>80%) occurred within the top 0–2 cm of sediment (Fig. S2). Generalized linear mixed-effects models indicated that both site and depth layer significantly affected all taxonomic and functional diversity indices (Table 4 & Tables S5–S11), with the exception of d, for which site was significant but depth layer was not (Table S8).

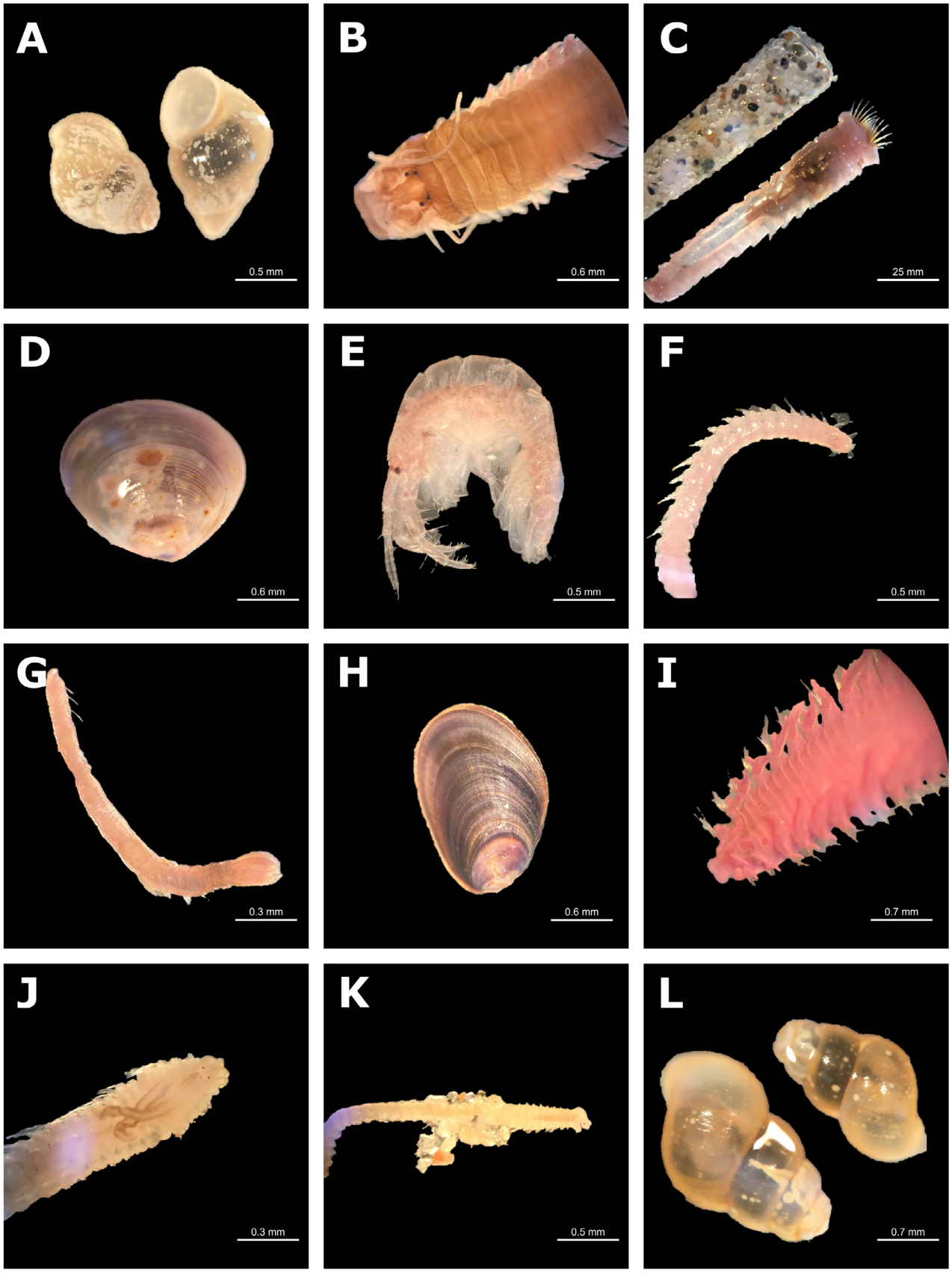

Fig. 2. Subset of macrofauna identified from sediment cores: (A) *Ecrobia truncata*, (B) *Alitta* sp., (C) *Pectinaria gouldii*, (D) *Gemma gemma*, (E) *Monocorophium* sp., (F) *Microphthalmus* sp., (G) *Manayunkia aestuarina*, (H) Mytilidae indet., (I) *Marenzelleria viridis*, (J) *Polydora cornuta*, (K) *Pygospio elegans*, (L) Rissooidea indet

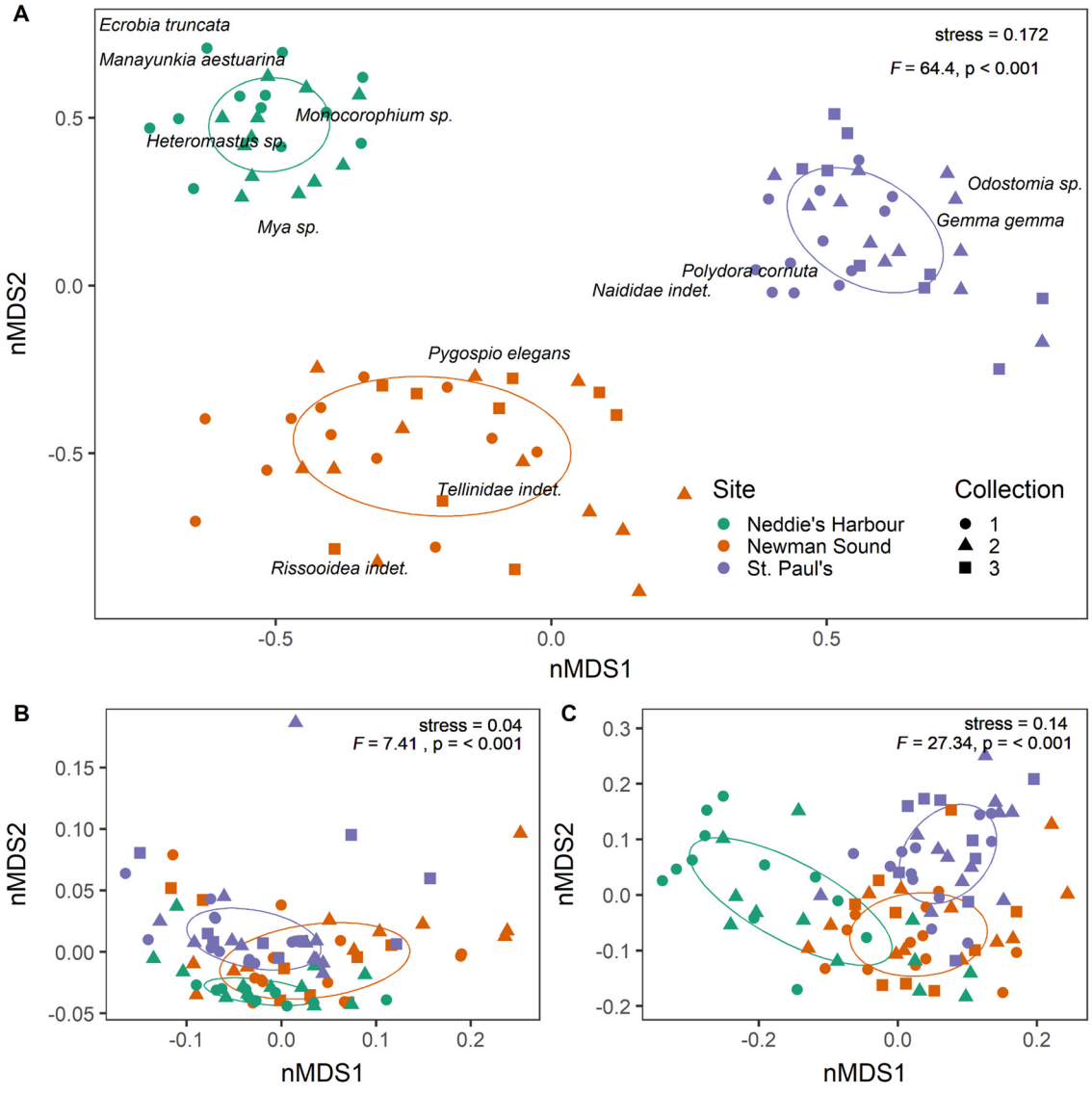

Fig. 3. Non-metric multidimensional scaling (nMDS) plots of (A) taxonomic composition, (B) taxonomic diversity indices, and (C) functional diversity indices at each site. Coloured symbols represent individual sediment cores (colour: site; shape: collection ID); collections 1, 2, and 3 represent the first, second, and third days of field sampling at each study site, respectively. Black text indicates individual taxa identified by SIMPER analysis. Taxa most strongly influence the taxonomic composition of cores to which they are most closely situated. Results of supporting PERMANOVAs (pseudo-*F* and p) are displayed in the top right corner of each plot

3.3.1. Taxonomic diversity

At all sites, *S* decreased with depth layer, but this change was most pronounced in NS (Table 4 & Table S5). Conversely, *J′* increased with depth, although the magnitude of this change differed among sites (Table 4 & Table S7). As a result, the highest values of *J′* in surface sediment (top 0–2 cm) in NH contrasted with the highest *J′* in deeper sediment (2–10 cm) in NS (Table 4). *H′* only differed between depth layers at NH, where it decreased with depth (Table S12). *H′* was also highest in NH, regardless of depth layer (Table 4 & Table S6). Finally, *d* did not differ significantly between depth layers but was significantly higher at NH than at the other 2 sites (Table 4 & Tables S8 & S12).

3.3.2. Functional diversity

In addition to significant depth layer and site differences, the site × depth layer interaction differed

Table 3. Taxa that contributed most to the dissimilarity between communities, presented in descending order according to their relative contributions. SIMPER analyses were performed between each pair of sites, and taxa that appeared at least twice in the top 10 most influential taxa are included. Mean density (ind. m^{-2}, rounded to the nearest whole number) ± standard deviation is listed for each taxon for each site

Taxon	Neddie's Harbour	Newman Sound	St. Paul's
Monocorophium sp.	1294 ± 977	37 ± 85	131 ± 169
Naididae indet.	986 ± 739	1151 ± 940	2342 ± 1416
Pygospio elegans	1368 ± 658	2875 ± 2104	632 ± 372
Ecrobia truncata	284 ± 302	0	0
Gemma gemma	0	0	5707 ± 4849
Heteromastus sp.	219 ± 134	21 ± 49	2 ± 12
Manayunkia aestuarina	815 ± 459	0	0
Mya sp.	993 ± 412	118 ± 98	58 ± 88
Odostomia sp.	0	0	681 ± 817
Polydora cornuta	121 ± 112	97 ± 90	711 ± 361
Rissooidea indet.	0	198 ± 298	0
Tellinidae indet.	0	511 ± 310	64 ± 80

significantly for each functional diversity index (Tables S9–S11). At all sites, FRic decreased with sediment depth (Table 4 & Table S9). However, this change was negligible in NH, where mean FRic decreased by only 0.01 (Table 4 & Table S12). Examining the other 2 indices — FEve and FDiv — revealed contrasting patterns among sites (Table 4). In NH, both FEve and FDiv increased with depth (Table 4). However, at the other 2 sites, change in FEve was negligible between depth layers (Table 4 & Table S12). FDiv increased with depth at SP but decreased with depth in NS (Table 4).

Some patterns in CWM related significantly to sediment depth (Fig. 7). For example, macrofauna in the deeper (2–10 cm) sediment layer were slightly larger and more mobile than those in the top 0–2 cm, except in SP where organisms in both depth layers were similar in size and mobility (Fig. 7, Tables S13 & S14). Patterns in diet were similar between depth layers, with herbivory dominating organisms (60% or more) in all site × depth layer combinations (Table S13). However, the proportion of organisms belonging to each feeding guild differed between depth layers. Across all sites, suspension feeders were roughly twice as common in the 0–2 cm sediment layer compared to the deeper 2–10 cm layer (Table S13); sub-surface deposit feeders were more abundant in deeper sediment, although this difference was not significant in NS (Tables S13 & S14). Sediment reworking also differed notably between depth layers. Biodiffusors were most common in the 2–10 cm layer compared to the 0–2 cm layer (Table S13); this difference was most evident in SP where biodiffusors such as Naididae indet., *Odostomia* sp., and *Alitta* sp. accounted for 48.1% of all macrofauna in the deeper sediment, compared to 14% of fauna in the surface layer. Unsurprisingly, surficial modifiers were most common in surface (0–2 cm) sediment, although again, this difference was most pronounced in SP (Table S13). Finally, upward and downward conveyors were more common in the deeper sediment layers; at NH, they made up 33.3% of the deeper sediment community compared to 14.8% of the surface sediment community. SP was similar, with conveyors making up 25.2% of organisms in the 5–10 cm layer but only 7.5% in the top 0–2 cm of sediment. NS was the exception, characterized by the prevalence of upward and downward conveyors (~50%) at both sediment depths.

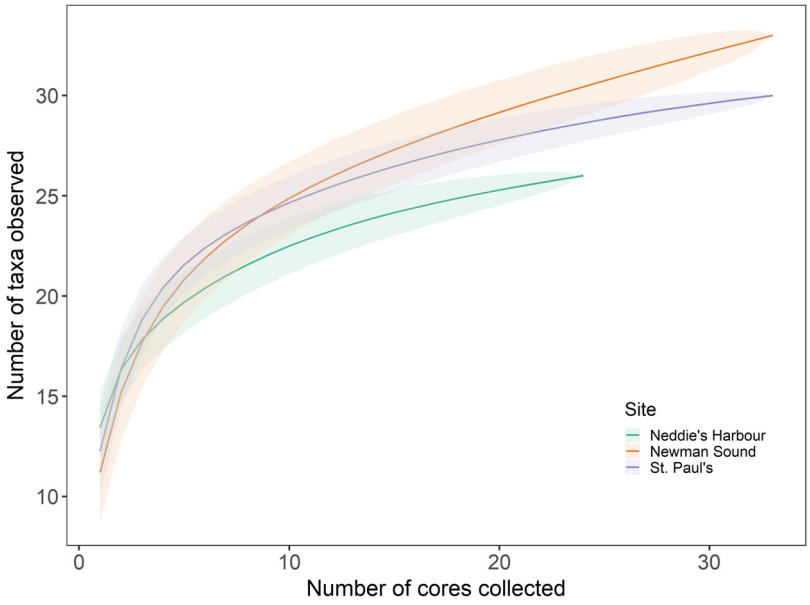

Fig. 4. Species accumulation curves for Neddie's Harbour, Newman Sound, and St. Paul's reflecting the number of cores collected at each study site and the total number of taxa observed in those cores. Solid lines: mean species richness; shaded areas: ±1 SD

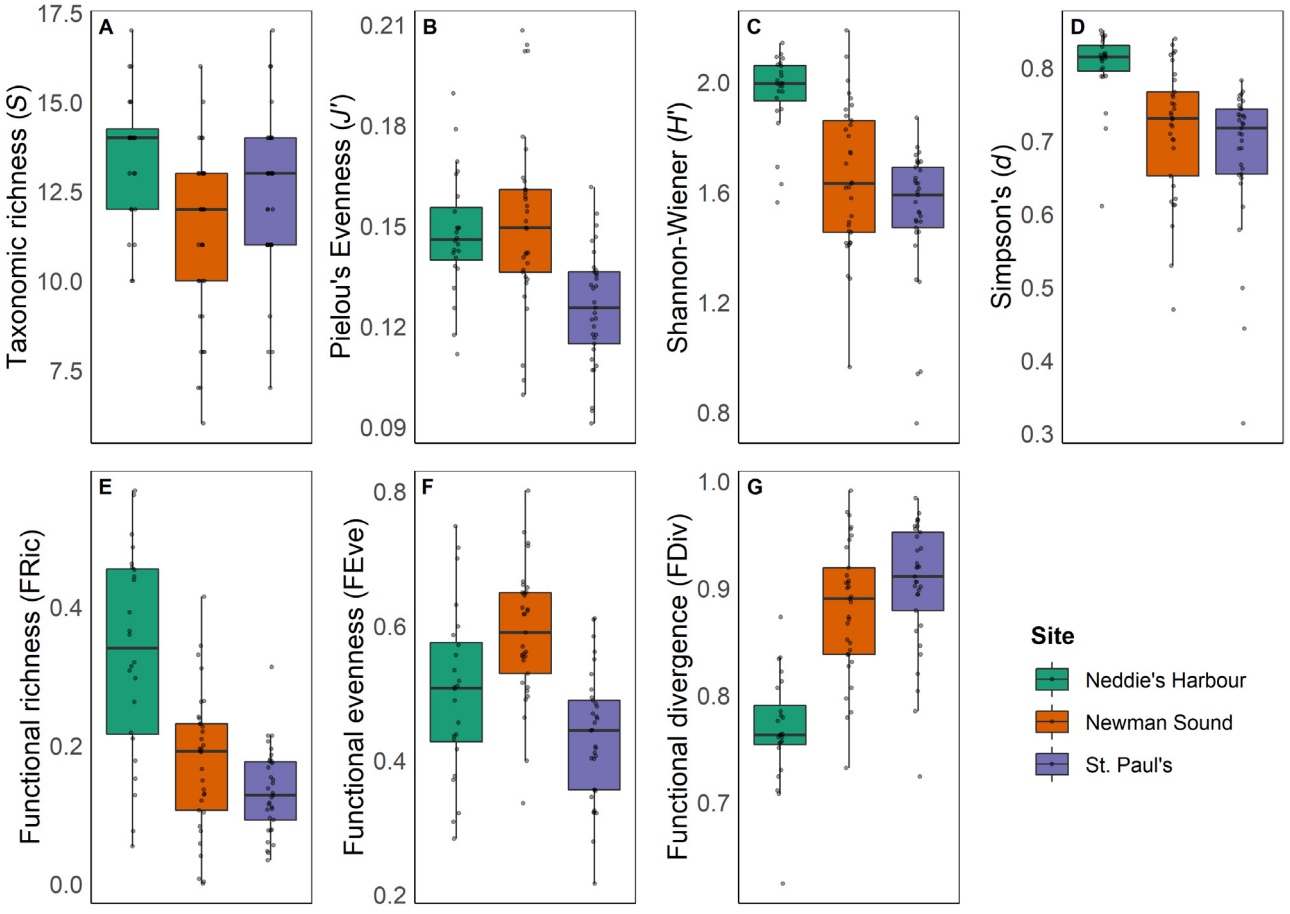

Fig. 5. Indices of (A–D) taxonomic diversity and (E–G) functional diversity from all cores at all study sites. Dark horizontal lines: median values; boxes: interquartile range (IQR); whiskers: range of data within ±1.5 times the IQR; small grey points: individual sediment cores collected at each site, including outliers

3.4. Relationship between taxonomic and functional diversity indices

Several patterns emerged in examining scatterplots and correlation coefficients to compare taxonomic and functional diversity indices across all sites (Fig. 8). H' and d were strongly and positively correlated (r = 0.95), noting a more left-skewed distribution for d than for H'. Positive relationships were also apparent between H' and FRic (r = 0.68), and S and FRic (r = 0.66). S and J' (r = −0.57) produced the strongest negative relationship. The weakest correlations occurred between FEve and d (r = −0.01) and between FEve and FDiv (r = −0.03).

3.5. Power analysis: effects of sample size on diversity assessment

After exploring results from the full set of samples, we conducted a post hoc power analysis to determine whether fewer replicates yield different conclusions. For many diversity indices, sample size influenced detection of statistical differences among sites, depth layers, and their interaction (Fig. S3, Table S15).

3.5.1. Detecting site-level effects

A sample size of 5 cores yielded high confidence that site affects H', d, and FDiv, with statistical powers exceeding 80% for all 3 indices (Fig. 9, Table S16). In other words, >80% of the 1000 model runs yielded a significant site effect for these community metrics ($p < 0.05$), even with a relatively small sample size. Other indices required larger sample sizes to reach the 80% power threshold, with FEve, FRic, and S requiring 10, 15, and 20 cores, respectively (Fig. 9, Table S16). Finally, in the case of J', even the maximum sample size of 20 cores resulted in a power of only 65.8% (Table S16). This finding suggests that assessing the effect of site on J' statistically would

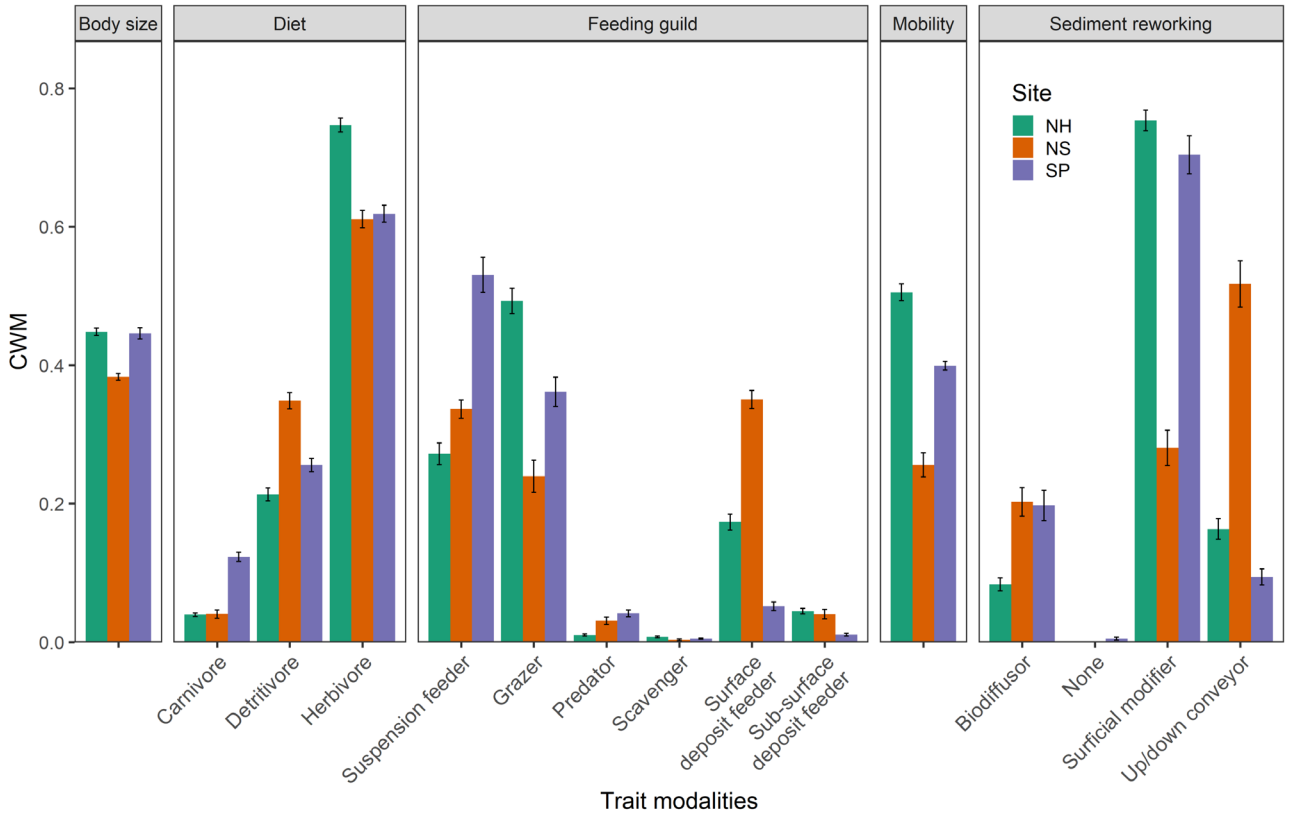

Fig. 6. Community-level weighted means (CWMs) for each functional trait, based on values calculated for each sediment core. The *x*-axes indicate modalities for each functional trait. For continuous variables (body size and mobility) *y*-axis indicates the mean value for that trait at each site. For categorical variables, *y*-axis represents the probability that an individual organism drawn at random from the community would exhibit that trait modality. Green, orange, and purple bars represent the macrofaunal communities at Neddie's Harbour (NH), Newman Sound (NS), and St. Paul's Bay (SP), respectively. Error bars: ±SE

require a sample size greater than that available in our study (Fig. 9).

For several indices, increasing sample size switched the detected effect of site on diversity from predominantly non-significant to predominantly significant (Fig. 9). *S* was the most prominent example of this phenomenon. With 3 cores, 80.1% of model iterations resulted in a non-significant p-value for site, whereas 20 cores produced a significant effect 95.8% of the time (Fig. 9, Table S16). A similar pattern was observed for FEve and FRic (Fig. 9).

Examining differences between individual pairs of sites added further insight. For most indices, a sample size of 3 cores was sufficient to separate the most dis-

Table 4. Diversity indices for each depth layer (cm) at each site: average (±SD) taxon richness (*S*), Shannon-Wiener index (*H′*), Simpson's diversity index (*d*), Pielou's evenness (*J′*), functional richness (FRic), functional evenness (FEve), and functional divergence (FDiv), based on measures for each sample. **Bold** values indicate significant differences ($p < 0.05$) between depth layers within a given site (see Table S12)

	Neddie's Harbour		Newman Sound		St. Paul's	
	0–2 cm	2–10 cm	0–2 cm	2–10 cm	0–2 cm	2–10 cm
S	**9.3 ± 1.7**	**6.6 ± 2.5**	**8.4 ± 1.9**	**4.3 ± 2.0**	**8.8 ± 2.4**	**5.4 ± 2.2**
H′	**1.8 ± 0.16**	**1.6 ± 0.35**	1.3 ± 0.33	1.2 ± 0.46	1.3 ± 0.23	1.2 ± 0.34
d	0.79 ± 0.043	0.74 ± 0.11	0.61 ± 0.14	0.62 ± 0.20	0.60 ± 0.11	0.61 ± 0.14
J′	**0.19 ± 0.027**	**0.26 ± 0.050**	**0.17 ± 0.041**	**0.28 ± 0.090**	**0.16 ± 0.056**	**0.24 ± 0.064**
FRic	0.14 ± 0.11	0.13 ± 0.11	**0.13 ± 0.064**	**0.042 ± 0.064**	**0.10 ± 0.065**	**0.030 ± 0.037**
FEve	**0.58 ± 0.17**	**0.72 ± 0.14**	0.59 ± 0.10	0.62 ± 0.21	0.50 ± 0.10	0.48 ± 0.15
FDiv	**0.75 ± 0.066**	**0.84 ± 0.079**	**0.90 ± 0.058**	**0.83 ± 0.10**	**0.86 ± 0.070**	**0.91 ± 0.063**

tinctive site from the other 2; however, differentiating between the 2 more similar sites (often NH and NS) required 5–20 cores, depending on the community metric in question (Fig. S3). The exceptions to this pattern were FEve and *S*, which required minimum sample sizes of 6 and 9 cores, respectively, to separate the most distinct site from the others (Fig. S3).

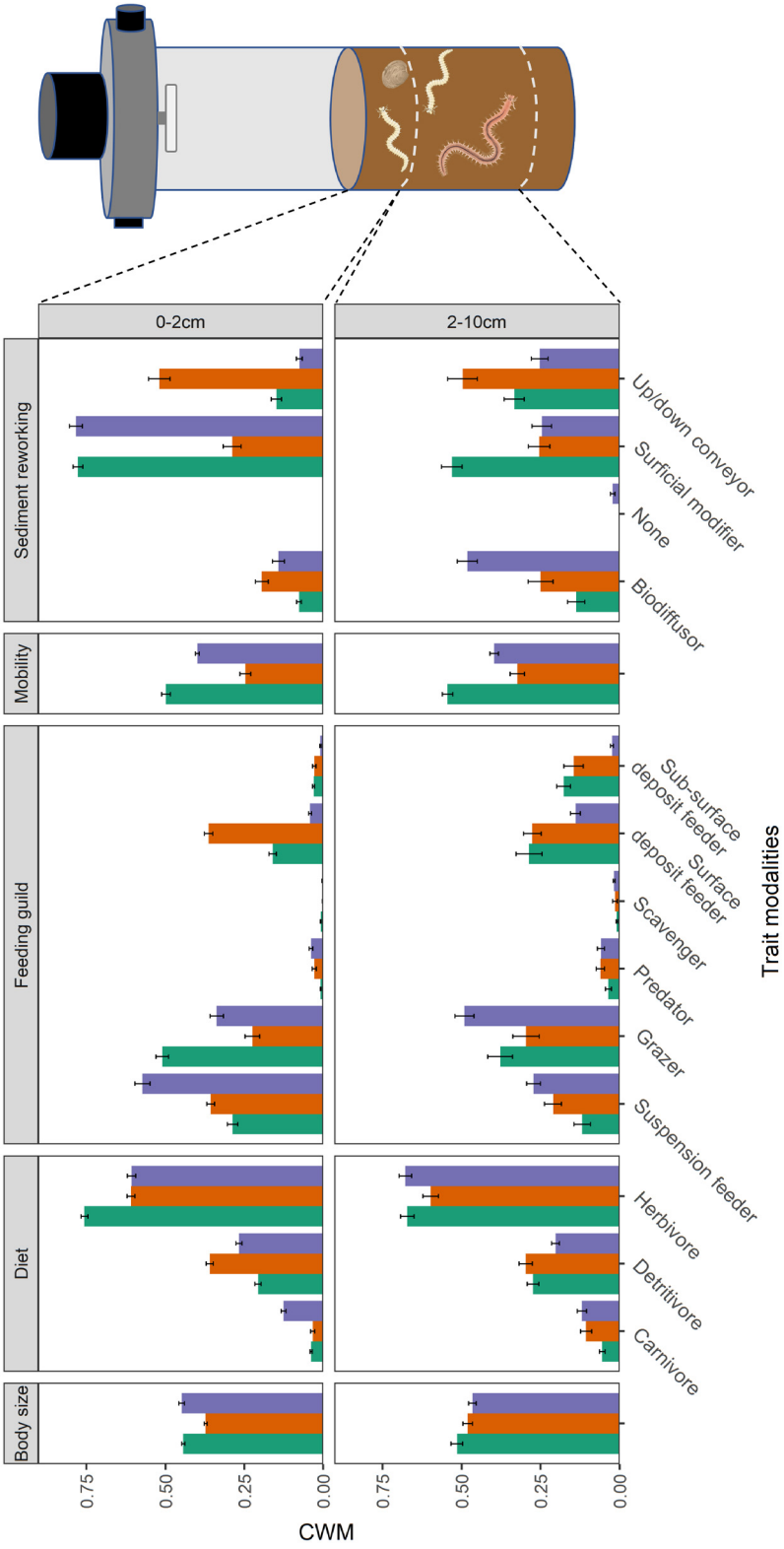

Fig. 7. Community-level weighted means (CWMs) for each functional trait, separated by depth layer: top row: macrofauna located in the upper 0–2 cm of sediment; bottom row: those found in the lower 2–10 cm of sediment. The *x*-axis indicates modalities for each functional trait. For continuous variables (body size and mobility), *y*-axis indicates the mean value for that trait at each site; for categorical variables, *y*-axis represents the probability that an individual organism drawn at random from the community would exhibit that trait modality. Green, orange, and purple bars represent Neddie's Harbour, Newman Sound, and St. Paul's, respectively. Error bars: ±1 SE around each CWM. Organism images courtesy of the IAN/UMCES Symbol and Image Libraries (Integration and Application Network; www.ian.umces.edu/media-library)

3.5.2. Detecting depth layer effects

For *S*, *J'*, and *d*, 3 cores were sufficient to achieve a power of >80% when assessing the effect of depth layer on diversity (Fig. 9, Table S16). FRic required only 5 cores to detect a statistically significant effect with high confidence. However, FDiv and *H'* both required ~20 cores to attain 80% power. For FDiv in particular, collection of 10 cores or fewer would have resulted in a relatively high probability (72.5%) of erroneously concluding that depth layer has no significant effect on this index (Fig. 9). For *H'* and FEve, the detected effect of depth layer on diversity also switched from non-significant to significant as sample size increased, although power for FEve did not surpass 80% for any sample size (Fig. 9).

Comparing depth layers within a given site typically required 3–10 cores to detect statistically significant differences in diversity indices, where they occurred (Fig. S4). However, when comparing diversity indices among sites for a given depth layer, the required sam-

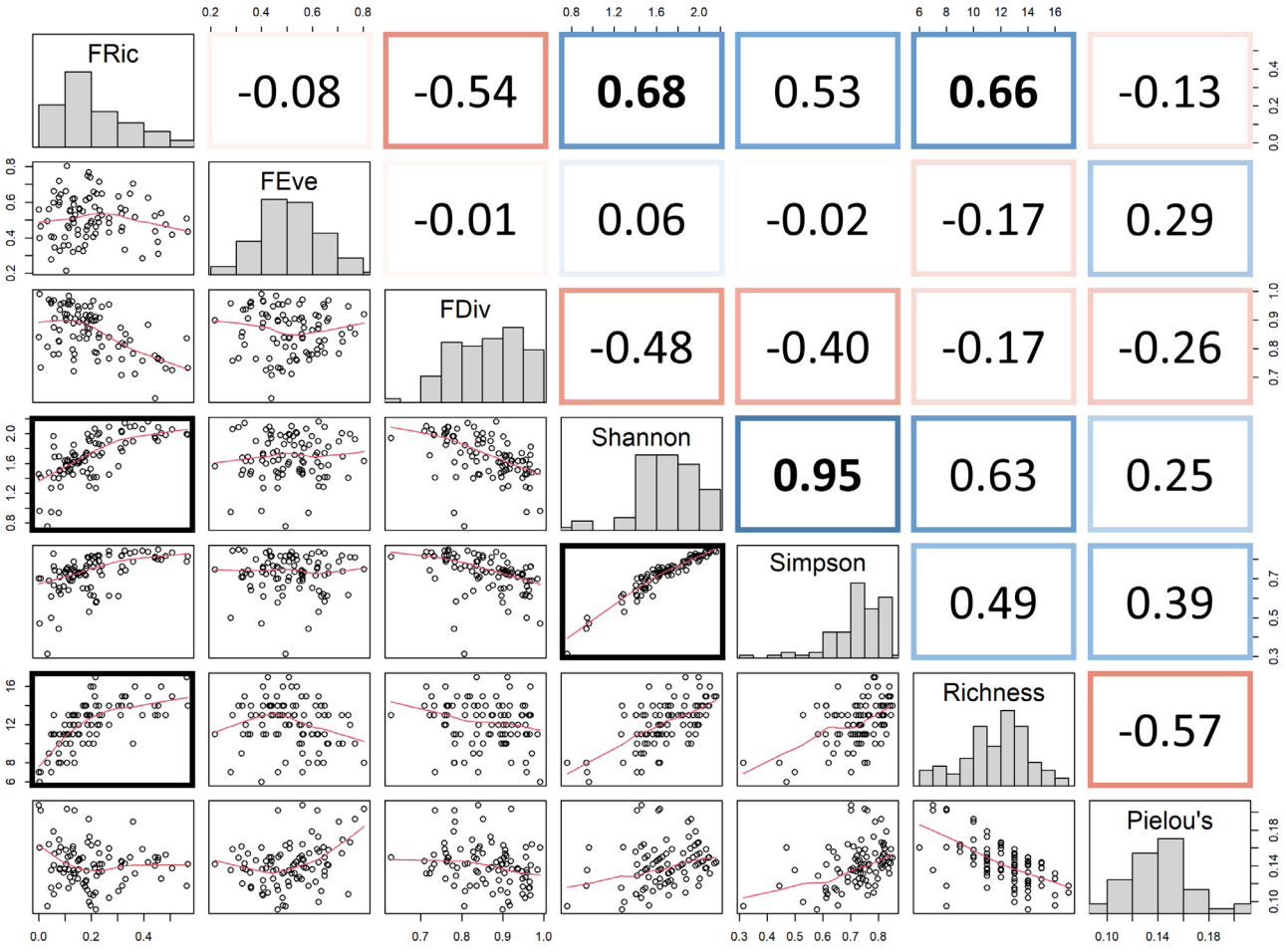

Fig. 8. Scatterplots and correlations between all taxonomic and functional diversity indices. Lower section is composed of scatterplots of each index regressed against the others; points represent individual sediment cores, red lines are the lowess smooth of the points. Upper corner lists correlation coefficients for each pair of diversity indices. Positive (blue) and negative (red) correlation coefficients are indicated by coloured borders. **Bold** values indicate high correlation coefficients (⩾0.7); the corresponding scatterplots for these highly correlated indices are outlined in **bold**

ple size varied from 3–20+ cores depending on the index and depth layer of interest (Fig. S4). For instance, in the 0–2 cm layer, a sample size of 3 cores was sufficient to differentiate NH from the other 2 sites with respect to H', d, and FDiv. However, 10 cores were required to separate FDiv for the other 2 sites, and >20 cores would be required to detect differences in H' and d, should they exist.

3.5.3. Detecting interactions

The interaction between site and depth layer had a statistically significant effect on both FDiv and FRic. However, the sample size required to detect this effect differed considerably between the 2 indices (Fig. 9). For FDiv, 5 cores were sufficient to yield a power of 86.8%, whereas FRic required 15 cores to achieve a similar result (Table S16). For FEve and J', 20 cores were insufficient to achieve a power of 80%, suggesting a larger sample size would be required to assess the interaction term for these indices. When modeling the full data set, we found no statistically significant interactions between site and depth layer for the other indices.

4. DISCUSSION

Our study emphasizes the importance of using multiple approaches to describe biological communities. Taxonomic and functional indices revealed complementary aspects of diversity, yielding valuable insights, particularly because the traits we examined reflect key elements of ecosystem functioning. We observed pronounced differences in community composition among all sites regardless of their geographic proximity, highlighting the potential role of

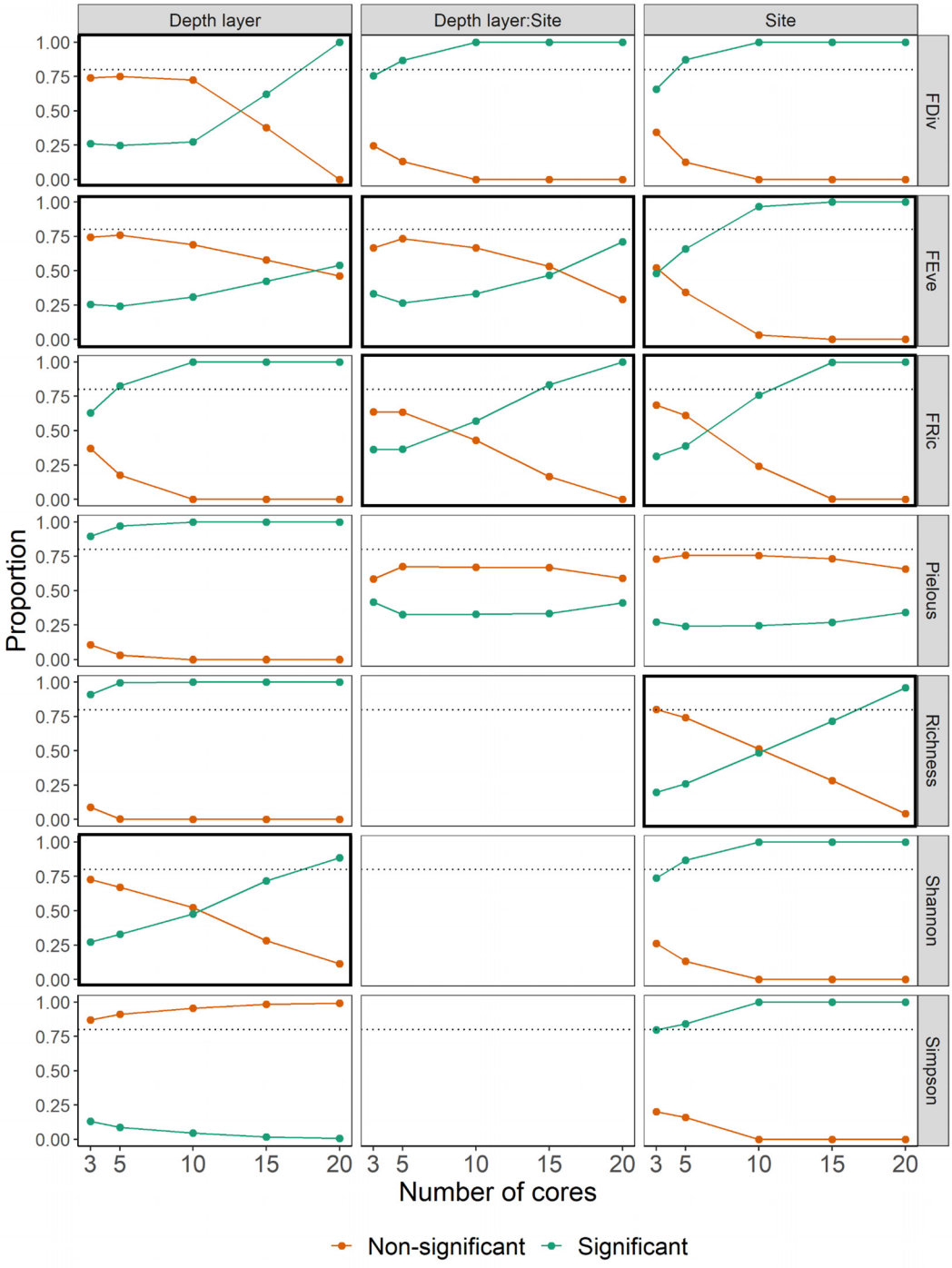

Fig. 9. Post hoc power analysis. Each model was fit 1000 times for each number of subset cores. Green and orange lines indicate the proportion of times a given model component was deemed significant ($p < 0.05$) or non-significant, respectively. Dashed line: the 'target' power of 80%. **Bold** boxes indicate index × model term combinations in which different sample sizes resulted in different statistical conclusions. Interaction terms did not have a significant effect on richness, Shannon-Wiener, or Simpson's indices, and are therefore not included in the power analysis

local and regional drivers in structuring infaunal communities. Relationships among commonly used diversity indices also varied; some metrics captured unique aspects of community composition, whereas others exhibited strong correlations but differed in their susceptibility to variation based on sample size. Finally, we found that typical sampling for coastal benthic studies (i.e. 3–5 cores site^{-1}) may be suit-

able for biodiversity assessment only in particular contexts, emphasizing the importance of clearly defining ecological questions prior to determining adequate sample sizes.

4.1. Regional context for macrofaunal community structure

Macrofaunal densities observed in our study fall within the range of previous reports for coastal and fjord sediments in Atlantic Canada (Ramey & Snelgrove 2003, Gerwing et al. 2015, Dreujou et al. 2020, Colvin & Snelgrove 2023, McGarrigle & Hunt 2024), although densities described in the Bonne Bay area nearly 20 yr ago (Quijón & Snelgrove 2005) were lower than those observed here. This discrepancy may partly be due to the different bottom depths sampled, given that we collected our cores intertidally rather than at a depth of 12–15 m. Additionally, we attribute the notably high mean density at SP, relative to most other sites in the region, to the high abundance of the bivalve *Gemma gemma*. Similar densities of *G. gemma* were reported in a coastal marine protected area in the southern Gulf of St. Lawrence, although the mechanisms contributing to this dominance were unclear (Lutz-Collins & Quijón 2014).

Assessments of shallow-water macroinfaunal diversity remain few in Eastern Canada, particularly with respect to functional metrics. However, we observed taxonomic and functional diversity values comparable to those recently reported for unvegetated sediments in Newman Sound on the east coast of Newfoundland (Colvin & Snelgrove 2023), with 2 exceptions. First, we observed lower J'; this pattern holds for other coastal studies in Newfoundland and the Gulf of St. Lawrence, which report similar H' values but higher J' than our study (Ramey & Snelgrove 2003, Dreujou et al. 2020). J' describes the extent to which abundances are uniformly distributed across taxa in a community (Pielou 1975). Low values are thus unsurprising for SP and NS, which were each dominated by a single taxon. In NH, 3 of the 25 taxa observed were highly abundant, collectively accounting for over 50% of macrofauna. This pattern resulted in lower J' than other studies in the region. Finally, we observed high site-level richness compared to other coastal and fjord areas, although this observation may reflect methodological differences (i.e. more extensive sampling) rather than ecological variation per se.

4.2. Potential drivers of community composition

Of the 53 taxa we observed, 16 were unique to NS and 20 were recorded only in our west coast study sites. Notably, many species that did not appear in NS (e.g. *G. gemma*, *Parexogone hebes*, *Marenzelleria viridis* etc.) have known distributions that include the Gulf of St. Lawrence but do not extend to eastern Newfoundland (Brunel et al. 1998). The interplay of several factors, including regional variation in species pools or environmental conditions, may drive differences in composition among sites (Menge & Olson 1990). For instance, seasonal sea ice cover typifies the Bonne Bay region (Steel et al. 1994) but not NS, resulting in contrasting conditions that may select for different communities. Indeed, species assemblages in the Bonne Bay area represent a transition between Arctic and sub-Arctic fauna (Brunel et al. 1998, Quijón & Snelgrove 2005). Local drivers may also play a role in structuring macrobenthic communities, as evidenced by significant differences in diversity among all study sites, even those in close geographic proximity.

4.2.1. Geological effects

The geological history of western Newfoundland may partially explain the differences between SP and NH. Bonne Bay is a compound fjord with a shallow (~14 m) sill separating NH (in the East Arm) from the rest of the bay (Richards & deYoung 2004). Sills can influence ecological conditions, including availability of organic matter (Aure & Stigebrandt 1989), water mass characteristics such as dissolved oxygen (Holte et al. 2005), and larval dispersal (e.g. Molinet et al. 2006), all of which may affect macrofaunal abundance and distribution. Additionally, lower sea levels historically isolated several fjords in the area from other water bodies (Butler et al. 1996); the East Arm of Bonne Bay was separated from the Gulf of St. Lawrence, potentially promoting variation in faunal assemblages across relatively small spatial scales (i.e. km). The East Arm now supports a mixture of Arctic and temperate species (Butler et al. 1996), contributing to the high taxonomic diversity (S, H', and d) and FRic in NH. By contrast, SP opens directly into the Gulf of St. Lawrence, experiences seasonal sea ice cover, and contains similar fauna to the Gulf itself (Brunel et al. 1998). Finally, 24 freshwater tributaries feed into St. Paul's inlet, potentially supporting the dominance of *G. gemma*, which tolerates estuarine conditions (Kennedy & Mihursky 1971, O'Sullivan 1976).

4.2.2. Sediment grain size

Sediment grain size may also influence the local dominance of certain functional groups. At SP, the sandiest site with the lowest proportions of mud, suspension feeders dominated. By contrast, relatively few suspension feeders were observed in NH, which contains the most mud, whilst NS contains intermediate proportions of both mud and suspension feeders. This pattern may reflect differences in hydrodynamic conditions; areas of reduced water flow are unfavourable for passive suspension feeding (Hentschel & Larson 2005) but allow mud to accumulate in surface sediment. Fine-grained sediments may also clog the filtering structures of some suspension feeders, reducing the fitness of this feeding mode in muddy environments (Shimeta & Jumars 1991).

4.2.3. Food availability

Of the 3 sites, NH contained the highest food quantity (%C), which is generally associated with increased faunal densities (Campanyà-Llovet et al. 2017); however, food quality was relatively low. Given the high C:N ratios in woody tissues relative to organic matter of marine origin (Meyers 1994, Hedges & Oades 1997), lower quality terrestrial inputs likely dominate organic matter in NH. Indeed, we observed woody material in surface sediment at this site (M. E. Clinton pers. obs.). Inclusion of stable isotope analysis in future studies could add clarity on this point. Although NH did not contain the highest overall macrofaunal density, the community exhibited higher S and J' than SP or NS. High food quantity may thus be supporting the co-dominance of several highly abundant taxa at this site.

By contrast, food quantity was lower at SP and NS but quality was higher, with C:N ratios suggesting proportionally greater inputs of marine food sources compared to NH. At SP, intense filter feeding may prevent the accumulation of organic matter in the sediment, explaining the low food quantities observed. Indeed, *G. gemma* often nearly or completely exploits its food resources when present in high abundances (Sanders et al. 1962). However, the relatively high chl *a*:Phaeo ratio at SP indicates fresh phytodetritus (Le Guitton et al. 2015) and suggests that benthic primary production (e.g. by diatoms) may be contributing to the high chl *a* at this site (Glud et al. 2002, Cox et al. 2020). The lower chl *a* in NS may be due to the prevalence of upward and downward conveyors, which often contribute to rapid subduction of organic matter as observed in other intertidal systems (e.g. Middelburg et al. 2000).

4.3. Contrasting effects of depth layer on infaunal diversity

Our study aligns with previous findings that sediment depth affects macroinfaunal abundance and diversity (Witte 2000, Celentano et al. 2019), but the direction and magnitude of these relationships varied among sites for several indices. H' differed between depth layers only at NH, where it was highest in the upper sediment layer (0–2 cm). This pattern is particularly interesting because d, which correlated strongly with H', exhibited no discernable relationship with depth layer at any site. Although both indices incorporate elements of taxonomic richness and evenness, H' places greater weight on richness whilst d is more heavily influenced by evenness (DeJong 1975). As such, the relative changes in richness and evenness between depth layers at NH may have been mathematically balanced for d but detectable for H', emphasizing the importance of incorporating multiple approaches into diversity assessments, even for correlated indices.

With respect to functional diversity, FDiv remained relatively consistent between depth layers at SP and NS but was notably low in the 0–2 cm depth layer at NH, suggesting little niche differentiation and thus higher potential for direct competition among dominant species in the surface sediment at this site (Mason et al. 2005, Mouchet et al. 2010, Dimitriadis et al. 2012). This pattern may be the result of the high quantity of terrestrially derived food at NH selecting for similar diets among abundant taxa, as evidenced by the high proportion of herbivores in the 0–2 cm layer.

FEve was also high in the 2–10 cm sediment layer at NH relative to the surface sediment, although this index did not appear to differ between depth layers at the other sites. High FEve indicates evenly distributed abundances within trait space (Villéger et al. 2008) and therefore similar abundances, or more regular functional distances, among species. The high FEve in the lower depth layer may simply reflect the relatively small proportion of fauna present in the 2–10 cm layer, consequently reducing possible variation in abundances among taxa. Alternatively, higher FEve may reflect more evenly distributed feeding traits in the 2–10 cm layer. Suspension feeders, grazers, and surface and sub-surface deposit feeders all occurred in relatively high proportions in the 2–10 cm sediment at this site, suggesting that deeper

infauna employ diverse strategies to exploit organic matter. Indeed, food quantity decreased with sediment depth at NH, potentially increasing interspecific competition and leading to greater variety and evenness of feeding traits in the deep sediment layer.

4.4. A multi-pronged approach to assessing diversity

Our results suggest that considering taxonomic and functional indices in tandem reveals complementary aspects of community structure, but that assessing any particular metric alone does not necessarily provide insight into other aspects of diversity, as illustrated by the range of correlation coefficients we observed between indices. Some indices (e.g. FEve and FDiv) were independent of all others considered, capturing unique aspects of community composition. Others, such as H' and d, correlated strongly but differed in susceptibility to variation based on sample size. Finally, CWMs revealed nuances that diversity indices alone miss. For example, CWMs captured the dominance of surficial modifiers at both west coast sites (NH and SP), while highlighting the abundance of upward and downward conveyors in NS — traits that serve as useful indicators of ecosystem functioning (Queirós et al. 2013).

To evaluate the effectiveness of conservation initiatives or assess impacts of anthropogenic activities using biodiversity data, many studies select simple indices such as S or abundance of key species (e.g. Lipej et al. 2003, Lester & Halpern 2008, Lewis et al. 2014). Acknowledging that conservation efforts may sometimes prioritize biodiversity per se, functional diversity metrics generally predict ecosystem functioning better than taxonomic diversity alone (Tilman et al. 1997, Bremner et al. 2006, Mouillot et al. 2011, Lefcheck & Duffy 2015, Lam-Gordillo et al. 2020). However, even functional diversity metrics may not adequately capture ecosystem status and trends when examined in isolation. Considering several dimensions of diversity concurrently may therefore yield a more comprehensive understanding of benthic structure and function and contribute to more accurate ecosystem assessments.

4.5. Implications of sample size for interpretation of diversity assessments

Most studies of marine infaunal diversity assess relatively large areas (i.e. 100s of km), with a small number of replicate samples collected at each station or site (e.g. Biles et al. 2002, Thrush et al. 2003, Belley & Snelgrove 2016, Miatta & Snelgrove 2021, Bianchelli et al. 2022). Time or resource availability in the field (e.g. ship time, equipment, or personnel) often limit sample size, along with the costly and labour-intensive process of identifying large quantities of macrofauna after collection. These limitations contribute to the dearth of sedimentary diversity studies that specifically address sample size effects on statistical power (but see Mavrič et al. 2013, Forcino et al. 2015).

Our power analysis illustrates that under-sampling could impede reliable assessment of biodiversity trends in marine ecosystems. Even basic estimates of taxon richness based on too few replicates can mislead, potentially obscuring biological hotspots or leading to inaccurate conclusions regarding diversity patterns in benthic systems (Fig. 10). In fact, for many indices (e.g. S, H', FRic, FDiv), few replicates led to a high probability of drawing erroneous conclusions with respect to the relationship between diversity and site or depth layer. Despite these limitations, some indices were relatively insensitive to changes in sampling effort. For example, 3–5 cores produced statistically robust results for d, suggesting that this index may be an appropriate metric for resource-limited diversity studies.

In general, taxonomic assessment required lower sample sizes compared to functional diversity. However, we observed multiple exceptions to this pattern, and for most indices considered, assessing site versus depth layer effects required different levels of replication. For example, 3 cores were sufficient to detect depth layer effects for all taxonomic diversity indices except H'. Conversely, S and J' required 20+ cores to distinguish among sites, indicating that the level of replication required depends strongly on the ecological question posed.

Although our analysis focused on numbers of samples, adjusting other parameters may also affect statistical power. For studies in which sample size is necessarily low (e.g. limited sampling time), larger cores may be more appropriate. However, larger cores require greater investment in laboratory processing which may not be feasible for all research programs. Altering core size can also influence diversity within size classes. For example, the relatively small cores in our study likely limited capture of large or fast-moving macrofauna. Given that different size classes of benthic organisms exhibit characteristic variation on different spatial scales (Levin 1992, Silberberger et al. 2018), different core sizes would likely capture different distribution patterns, altering observed diversity metrics. As such, the size range of the organisms of interest merits consideration in relation to the size of the sampling unit when determining appropriate levels of replication for benthic diversity assessment.

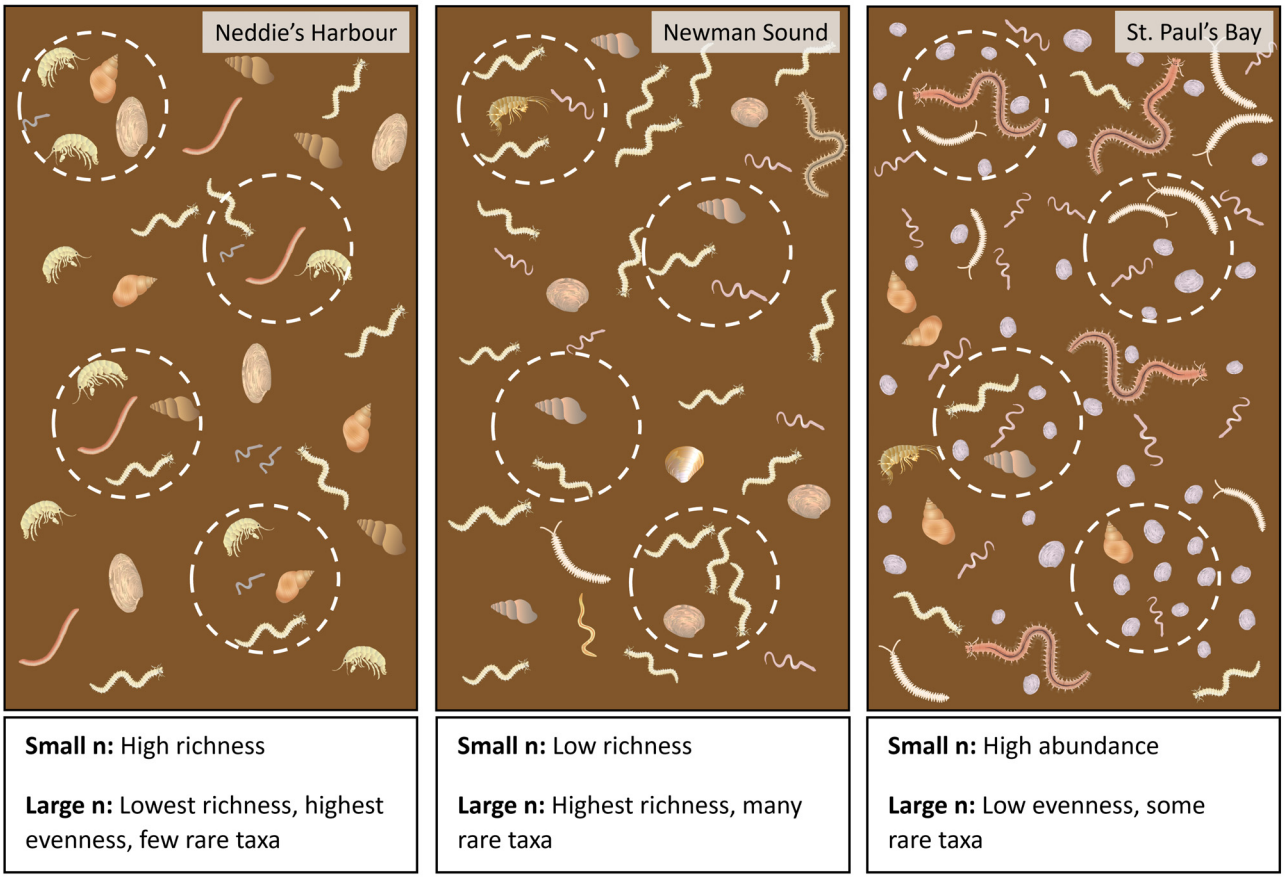

Fig. 10. Visual representation of community composition at each study site. Dashed white circles: potential sediment core collections. Differences in macrofaunal abundance, distribution (i.e. patchiness), overall richness, and number of rare species influence ecological conclusions that may be drawn based on varying sample sizes (n). Text in figure reports overall conclusions regarding each of the 3 study sites based on different levels of sampling effort. Organism images courtesy of the IAN/UMCES Symbol and Image Libraries (Integration and Application Network; ian.umces.edu/media-library)

Taxonomic resolution may also affect the level of replication required to assess diversity. Our study binned or omitted several taxa from the analyses to compare sites with differing levels of taxonomic uncertainty directly (see the Appendix). This step was necessary to reduce biases and standardize taxonomic detail, and likely had minimal effect on indices such as FRic, depending on the functional traits available at the genus or family level. However, availability of species-level information for all organisms may have yielded higher values for certain indices (e.g. *S*), potentially resulting in greater differences in community composition among sites (i.e. larger effect sizes). Given that larger effects are more easily detected, taxonomic resolution and required sample size are likely inversely proportional. This consideration is particularly relevant in remote or poorly accessible study areas, such as polar and sub-polar environments with limited available data.

Overall, our findings suggest a need for caution when interpreting patterns in benthic diversity studies with small sample sizes. However, depending on the aspects of diversity considered, few samples (i.e. 3–5 cores) may sometimes be sufficient to draw statistically robust conclusions, thus rendering coastal research or monitoring programs more resource-efficient. The importance of diversity as a predictor of ecosystem functioning (Tilman et al. 1997, Díaz & Cabido 2001, Hooper et al. 2005, Mouillot et al. 2011) punctuates the need to determine appropriate sample sizes for a range of marine systems in order to assess biodiversity status and trends and prioritize conservation resources in a rapidly changing global ocean.

5. SUMMARY

Accurate quantification of biodiversity is a central challenge of ecological research and is crucial to the

success of conservation and monitoring programs globally. Our findings highlight the value of combining approaches to assess multiple elements of diversity, including both taxonomic and functional composition of communities. Although our study was observational and did not allow direct investigation of mechanistic links, our results also suggest several potentially important drivers of diversity, including geological history and small-scale food quality and quantity. Finally, we emphasize the importance of carefully selecting biodiversity metrics prior to designing sampling protocols for benthic diversity studies. Based on our results, we recommend a minimum sample size of 5 replicates (for push cores ~7 cm in diameter) for coastal sedimentary environments when assessing metrics such as *d*. However, additional samples are recommended when measuring other metrics of diversity. Future research is required to determine optimal sample sizes for a range of benthic environments, which will help minimize research costs while ensuring statistically robust conclusions regarding the status and trends of seafloor biodiversity.

Acknowledgements. This research was sponsored by Ocean Frontier Institute (Module E: Ecosystem Indicators), supported by the Canada First Research Excellence Fund. M.E.C. was supported by a National Sciences and Engineering Research Council of Canada (NSERC) C-GSM scholarship. We thank T. Colvin, R. Command, and V. Byrne for assistance with sediment collections, and Dr. R. Gregory (Fisheries and Oceans Canada), Dr. D. McIlroy, Dr. B. Hooper, and D. Rumbolt (Bonne Bay Marine Station) for providing field support and laboratory space. We also thank J. Wells, G. Van Biesen, and S. Jones (Memorial University of Newfoundland), as well as O. Brown (Natural Resources Canada) for their assistance with sediment organic matter and grain size analyses, and M. Miatta and R. Evans for assistance with photopigment analysis. We thank T. Macdonald and the team at Biologica Environmental Services for providing taxonomic expertise and macrofaunal reference collections. Finally, we thank Drs. H. Link, P. Renaud and E. Kristensen for helpful comments that greatly improved an earlier draft of the manuscript.

LITERATURE CITED

Aure J, Stigebrandt A (1989) On the influence of topographic factors upon the oxygen consumption rate in sill basins of fjords. Estuar Coast Shelf Sci 28:59–69

Baldrighi E, Lavaleye M, Aliani S, Conversi A, Manini E (2014) Large spatial scale variability in bathyal macrobenthos abundance, biomass, α- and β-diversity along the Mediterranean continental margin. PLOS ONE 9: e107261

Bates D, Maechler M, Bolker BM, Walker S (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67: 1–48

Belley R, Snelgrove PVR (2016) Relative contributions of biodiversity and environment to benthic ecosystem functioning. Front Mar Sci 3:242

Bianchelli S, Martire ML, Pola L, Gambi C, Fanelli E, Danovaro R, Corinaldesi C (2022) Impact of hypersaline brines on benthic meio- and macrofaunal assemblages: a comparison from two desalination plants of the Mediterranean Sea. Desalination 532:115756

Biles CL, Paterson DM, Ford RB, Solan M, Raffaelli DG (2002) Bioturbation, ecosystem functioning and community structure. Hydrol Earth Syst Sci 6:999–1005

Bremner J, Rogers SI, Frid CLJ (2003) Assessing functional diversity in marine benthic ecosystems: a comparison of approaches. Mar Ecol Prog Ser 254:11–25

Bremner J, Rogers S, Frid C (2006) Methods for describing ecological functioning of marine benthic assemblages using biological traits analysis (BTA). Ecol Indic 6:609–622

Brooks ME, Kristensen K, van Benthem KJ, Magnusson A and others (2017) glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. R J 9:378–400

Brunel P, Bossé L, Lamarche G (1998) Catalogue of the marine invertebrates of the Estuary and Gulf of Saint Lawrence. Can Spec Publ Fish Aquat Sci 126:1–405

Butler M, Chaisson RD, Daury RW, Dean S, Dietz SB, MacKinnon N, Steele J (1996) By the sea: a guide to the coastal zone of Atlantic Canada. Department of Fisheries and Oceans, Moncton

Campanyà-Llovet N, Snelgrove PVR, Parrish CC (2017) Rethinking the importance of food quality in marine benthic food webs. Prog Oceanogr 156:240–251

Cardinale BJ, Duffy JE, Gonzalez A, Hooper DU and others (2012) Biodiversity loss and its impact on humanity. Nature 486:59–67

Celentano E, Lercari D, Maneiro P, Rodríguez P and others (2019) The forgotten dimension in sandy beach ecology: vertical distribution of the macrofauna and its environment. Estuar Coast Shelf Sci 217:165–172

Collinge SK (2001) Spatial ecology and biological conservation. Biol Conserv 100:1–2

Colvin TJ, Snelgrove PVR (2023) Effects of disturbance on macrofaunal biodiversity–ecosystem functioning relationships in seagrass habitats. Mar Ecol 44:e12753

Convention on Biological Diversity (2022) Kunming-Montreal Global Biodiversity Framework. United Nations Environment Programme, Montreal

Cornwell WK, Schwilk DW, Ackerly DD (2006) A trait-based test for habitat filtering: convex hull volume. Ecology 87: 1465–1471

Cox TE, Cebrian J, Tabor M, West L, Krause JW (2020) Do diatoms dominate benthic production in shallow systems? A case study from a mixed seagrass bed. Limnol Oceanogr Lett 5:425–434

Danovaro R (2010) Methods for the study of deep-sea sediments, their functioning and biodiversity. CRC Press, Boca Raton, FL

Degen R, Faulwetter S (2019) The Arctic Traits Database — a repository of Arctic benthic invertebrate traits. Earth Syst Sci Data 11:301–322

Degen R, Aune M, Bluhm BA, Cassidy C and others (2018) Trait-based approaches in rapidly changing ecosystems: a roadmap to the future polar oceans. Ecol Indic 91:722–736

DeJong TM (1975) A comparison of three diversity indices based on their components of richness and evenness. Oikos 26:222–227

Díaz S, Cabido M (2001) Vive la différence: plant functional diversity matters to ecosystem processes. Trends Ecol Evol 16:646–655

Dimitriadis C, Evagelopoulos A, Koutsoubas D (2012) Functional diversity and redundancy of soft bottom communities in brackish waters areas: local vs regional effects. J Exp Mar Biol Ecol 426-427:53–59

Dray S, Dufour A (2007) The ade4 package: implementing the duality diagram for ecologists. J Stat Softw 22:1–20

Dreujou E, McKindsey C, Grant C, Tréau De Coeli L, St-Louis R, Archambault P (2020) Biodiversity and habitat assessment of coastal benthic communities in a sub-Arctic industrial harbor area. Water 12:2424

Ellingsen KE (2002) Soft-sediment benthic biodiversity on the continental shelf in relation to environmental variability. Mar Ecol Prog Ser 232:15–27

Fauchald K, Jumars PA (1979) The diet of worms: a study of polychaete feeding guilds. Oceanogr Mar Biol Annu Rev 17:193–284

Faulwetter S, Markantonatou V, Pavloudi C, Papageorgiou N and others (2014) Polytraits: a database on biological traits of marine polychaetes. Biodivers Data J 2:e1024

Ferrari S, Cribari-Neto F (2004) Beta regression for modelling rates and proportions. J Appl Stat 31:799–815

Fofonoff PW, Ruiz GM, Steves B, Simkanin C, Carlton JT (2018) National exotic marine and estuarine species information system. https://invasions.si.edu/nemesis/

Forcino FL, Leighton LR, Twerdy P, Cahill JF (2015) Reexamining sample size requirements for multivariate, abundance-based community research: when resources are limited, the research does not have to be. PLOS ONE 10: e0128379

Gerwing TG, Gerwing AMA, Drolet D, Barbeau MA, Hamilton DJ (2015) Spatiotemporal variation in biotic and abiotic features of eight intertidal mudflats in the upper Bay of Fundy, Canada. Northeast Nat 22:1–44

Glud RN, Kühl M, Wenzhöfer F, Rysgaard S (2002) Benthic diatoms of a high Arctic fjord (Young Sound, NE Greenland): importance for ecosystem primary production. Mar Ecol Prog Ser 238:15–29

Godbold JA, Solan M (2009) Relative importance of biodiversity and the abiotic environment in mediating an ecosystem process. Mar Ecol Prog Ser 396:273–282

Halekoh U, Højsgaard S (2014) A Kenward-Roger approximation and parametric bootstrap methods for tests in linear mixed models — the R package pbkrtest. J Stat Soft 59:1–32

Hartig F (2022) DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression models. R package v. 0.4.6, https://CRAN.R-project.org/package=DHARMa

Hedges JI, Oades JM (1997) Comparative organic geochemistries of soils and marine sediments. Org Geochem 27: 319–361

Hentschel BT, Larson AA (2005) Growth rates of interface-feeding polychaetes: combined effects of flow speed and suspended food concentration. Mar Ecol Prog Ser 293: 119–129

Holte B, Oug E, Dahle S (2005) Soft-bottom fauna and oxygen minima in sub-Arctic north Norwegian marine sill basins. Mar Biol Res 1:85–96

Hooper DU, Chapin FS, Ewel JJ, Hector A and others (2005) Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. Ecol Monogr 75: 3–35

Kennedy VS, Mihursky JA (1971) Upper temperature tolerances of some estuarine bivalves. Chesapeake Sci 12: 193–204

Kieschnick R, McCullough BD (2003) Regression analysis of variates observed on (0, 1): percentages, proportions and fractions. Stat Model 3:193–213

Laliberté E, Legendre P (2010) A distance-based framework for measuring functional diversity from multiple traits. Ecology 91:299–305

Laliberté E, Legendre P, Shipley B (2014) FD: measuring functional diversity from multiple traits, and other tools for functional ecology. R package v. 1.0-12, https://CRAN.R-project.org/web/packages/FD/index.html

Lam-Gordillo O, Baring R, Dittmann S (2020) Ecosystem functioning and functional approaches on marine macrobenthic fauna: a research synthesis towards a global consensus. Ecol Indic 115:106379

Lavorel S, Grigulis K, McIntyre S, Williams NSG and others (2008) Assessing functional diversity in the field — methodology matters! Funct Ecol 22:134–147

Le Guitton M, Soetaert K, Damsté JSS, Middelburg JJ (2015) Biogeochemical consequences of vertical and lateral transport of particulate organic matter in the southern North Sea: a multiproxy approach. Estuar Coast Shelf Sci 165:117–127

Lefcheck JS, Duffy JE (2015) Multitrophic functional diversity predicts ecosystem functioning in experimental assemblages of estuarine consumers. Ecology 96:2973–2983

Lefcheck JS, Bastazini VAG, Griffin JN (2015) Choosing and using multiple traits in functional diversity research. Environ Conserv 42:104–107

Legendre P, Fortin MJ (1989) Spatial pattern and ecological analysis. Vegetatio 80:107–138

Lester SE, Halpern BS (2008) Biological responses in marine no-take reserves versus partially protected areas. Mar Ecol Prog Ser 367:49–56

Levin SA (1992) The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. Ecology 73: 1943–1967

Lewis S, Ramirez-Luna V, Templeman N, Simpson MR and others (2014) A framework for the identification of monitoring indicators protocols and strategies for the proposed Laurentian Channel Marine Protected Area (MPA). Can Sci Advis Sec Res Doc 2014/093

Lipej L, Bonaca MO, Šiško M (2003) Coastal fish diversity in three Marine Protected Areas and one unprotected area in the Gulf of Trieste (Northern Adriatic). Mar Ecol 24: 259–273

Lipsey MW (1990) Design sensitivity: statistical power for experimental research, 9th edn. Sage Publications, Newbury Park, CA

Lutz-Collins V, Quijón PA (2014) Animal–sediment relationships in an Atlantic Canada marine protected area: richness, composition and abundance in relation to sediment food indicators. Mar Biol Res 10:577–588

Macdonald TA, Burd BJ, Macdonald VI, van Roodselaar A (2010) Taxonomic and feeding guild classification for the marine benthic macroinvertebrates of the Strait of Georgia, British Columbia. Can Tech Rep Fish Aquat Sci 2874: iv + 63 p

Magurran AE (1988) Ecological diversity and its measurement. Springer Science & Business Media, London

Mason NWH, Mouillot D, Lee WG, Wilson JB (2005) Functional richness, functional evenness and functional divergence: the primary components of functional diversity. Oikos 111:112–118

Mavrič B, Urbanič G, Lipej L, Simboura N (2013) Influence of sample size on ecological status assessment using marine benthic invertebrate-based indices. Mar Ecol 34:72–79

McGarrigle SA, Hunt HL (2024) Infaunal invertebrate community relationships to water column and sediment abiotic conditions. Mar Biol 171:3

Menge BA, Olson AM (1990) Role of scale and environmental factors in regulation of community structure. Trends Ecol Evol 5:52–57

Meyers PA (1994) Preservation of elemental and isotopic source identification of sedimentary organic matter. Chem Geol 114:289–302

Miatta M, Snelgrove PVR (2021) Benthic nutrient fluxes in deep-sea sediments within the Laurentian Channel MPA (eastern Canada): the relative roles of macrofauna, environment, and sea pen octocorals. Deep Sea Res I 178: 103655

Miatta M, Bates AE, Snelgrove PVR (2021) Incorporating biological traits into conservation strategies. Annu Rev Mar Sci 13:421–443

Middelburg JJ, Barranguet C, Boschker HTS, Herman PMJ, Moens T, Heip CHR (2000) The fate of intertidal microphytobenthos carbon: an *in situ* ^{13}C-labeling study. Limnol Oceanogr 45:1224–1234

Molinet C, Valle-Levinson A, Moreno CA, Cáceres M, Bello M, Castillo M (2006) Effects of sill processes on the distribution of epineustonic competent larvae in a stratified system of Southern Chile. Mar Ecol Prog Ser 324:95–104

Mouchet MA, Villéger S, Mason NWH, Mouillot D (2010) Functional diversity measures: an overview of their redundancy and their ability to discriminate community assembly rules. Funct Ecol 24:867–876

Mouillot D, Villéger S, Scherer-Lorenzen M, Mason NWH (2011) Functional structure of biological communities predicts ecosystem multifunctionality. PLOS ONE 6:e17476

Mouillot D, Graham NAJ, Villéger S, Mason NWH, Bellwood DR (2013) A functional approach reveals community responses to disturbances. Trends Ecol Evol 28: 167–177

Naeem S, Chazdon R, Duffy JE, Prager C, Worm B (2016) Biodiversity and human well-being: an essential link for sustainable development. Proc R Soc B 283:20162091

O'Sullivan W (1976) Hydrological inventory of Gros Morne National Park, Newfoundland Canada. Inland Water Directorate, Water Resources Branch, Halifax

Oksanen J, Blanchet FG, Friendly M, Kindt R and others (2020) vegan: community ecology package. R package v. 2.5-7, https://CRAN.R-project.org/package=vegan

Patrício J, Adão H, Neto JM, Alves AS, Traunspurger W, Marques JC (2012) Do nematode and macrofauna assemblages provide similar ecological assessment information? Ecol Indic 14:124–137

Petchey OL, Gaston KJ (2002) Functional diversity (FD), species richness and community composition. Ecol Lett 5:402–411

Pielou EC (1975) Ecological diversity. Wiley, New York, NY

Pinto R, Patrício J, Baeta A, Fath BD, Neto JM, Marques JC (2009) Review and evaluation of estuarine biotic indices to assess benthic condition. Ecol Indic 9:1–25

Polytraits Team (2022) Polytraits: a database on biological traits of polychaetes. http://polytraits.lifewatchgreece.eu/

Purvis A, Hector A (2000) Getting the measure of biodiversity. Nature 405:212–219

Queirós AM, Birchenough SNR, Bremner J, Godbold JA and others (2013) A bioturbation classification of European marine infaunal invertebrates. Ecol Evol 3: 3958–3985

Quijón PA, Snelgrove PVR (2005) Polychaete assemblages of a sub-Arctic Newfoundland fjord: habitat, distribution, and identification. Polar Biol 28:495–505

R Core Team (2021) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, https://www.R-project.org/

Ramey PA, Snelgrove PVR (2003) Spatial patterns in sedimentary macrofaunal communities on the south coast of Newfoundland in relation to surface oceanography and sediment characteristics. Mar Ecol Prog Ser 262: 215–227

Richards C, deYoung B (2004) Analysis of physical oceanographic data from Bonne Bay, September 2002–September 2004. Memorial University of Newfoundland, St. John's

Ricotta C, Moretti M (2011) CWM and Rao's quadratic diversity: a unified framework for functional ecology. Oecologia 167:181–188

Rietkerk M, van de Koppel J (2008) Regular pattern formation in real ecosystems. Trends Ecol Evol 23:169–175

Sanders HL, Goudsmit EM, Mills EL, Hampson GE (1962) A study of the intertidal fauna of Barnstable Harbor, Massachusetts. Limnol Oceanogr 7:63–79

Seed R (1996) Patterns of biodiversity in the macro-invertebrate fauna associated with mussel patches on rocky shores. J Mar Biol Assoc UK 76:203–210

Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27:379–423

Shimeta J, Jumars PA (1991) Physical mechanisms and rates of particle capture by suspension-feeders. Oceanogr Mar Biol Annu Rev 29:191–257

Silberberger MJ, Renaud PE, Buhl-Mortensen L, Ellingsen IH, Reiss H (2018) Spatial patterns in sub-Arctic benthos: multiscale analysis reveals structural differences between community components. Ecol Monogr 89:e01325

Simpson EH (1949) Measurement of diversity. Nature 163: 688

Steel A, Clark JI, Morin P (1994) A comparison of pressuremeter test results in sea ice. Can Geotech J 31:254–260

Steidl RJ, Thomas L (2001) Power analysis and experimental design. In: Scheiner SM, Gurevitch J (eds) Design and analysis of ecological experiments. Oxford University Press, New York, NY, p 14–36

Swift DJ (1993) The macrobenthic infauna off Sellafield (North-Eastern Irish Sea) with special reference to bioturbation. J Mar Biol Assoc UK 73:143–162

Thrush SF, Hewitt JE, Norkko A, Nicholls PE, Funnell GA, Ellis JI (2003) Habitat change in estuaries: predicting broad-scale responses of intertidal macrofauna to sediment mud content. Mar Ecol Prog Ser 263:101–112

Tilman D (2001) Functional diversity. In: Levin SA (ed) Encyclopedia of biodiversity. Academic Press, Cambridge, p 109–120

Tilman D, Knops J, Wedin D, Reich P, Ritchie M, Siemann E (1997) The influence of functional diversity and composition on ecosystem processes. Science 277:1300–1302

Villéger S, Mason NWH, Mouillot D (2008) New multidimensional functional diversity indices for a multifaceted framework in functional ecology. Ecology 89: 2290–2301

Violle C, Navas ML, Vile D, Kazakou E, Fortunel C, Hummel I, Garnier E (2007) Let the concept of trait be functional! Oikos 116:882–892
Wabnitz CCC, Andréfouët S, Muller-Karger FE (2010) Measuring progress toward global marine conservation targets. Front Ecol Environ 8:124–129
Witte U (2000) Vertical distribution of metazoan macrofauna within the sediment at four sites with contrasting food supply in the deep Arabian Sea. Deep Sea Res II 47: 2979–2997
WoRMS Editorial Board (2022) World Register of Marine Species. https://www.marinespecies.org/
Wu R (1982) Effects of taxonomic uncertainty on species diversity indices. Mar Environ Res 6:215-225

Appendix. Taxonomic resolution for calculation of diversity indices

When calculating diversity indices for each site, we made several adjustments to the taxonomic groupings to reduce error and biases. First, taxonomic resolution sometimes differed between sites for a given taxonomic unit. Where this occurred, we grouped taxa to the coarsest taxonomic level shared by all relevant sites. For example, the clam *Macoma balthica* was identified to species level in both St. Paul's Bay and Neddie's Harbour. However, in Newman Sound, it was only possible to identify *Macoma* clams to genus due to deterioration of key shell features. Therefore, all individuals of the genus *Macoma* across all sites were assigned the taxonomic label *Macoma* sp. for the purpose of calculating diversity indices.

In a few cases, we assigned a high taxonomic rank (low taxonomic resolution) to a solitary individual or to a small number (<5) of juveniles. In these situations, we omitted the individual or small group of juveniles when calculating diversity indices rather than grouping all taxonomic children to this coarse level. For example, one juvenile clam at Neddie's Harbour was identified as Macominae indet. We removed this individual from the diversity analysis, rather than labeling all *Macoma* sp. at all sites as Macominae indet. Where taxonomic resolution was identical between sites for a given taxonomic unit, we made no changes to taxonomic groupings.

Finally, previous research indicates that significant errors can occur when a taxonomic group with a large number of individuals is not differentiated into smaller taxonomic units and is instead treated as a single unit when calculating diversity indices (Wu 1982). We therefore removed 2 highly abundant (n > 1000) class-level taxonomic groups (Bivalvia indet. and Gastropoda indet.) from all sites for the purposes of calculating diversity indices. We decided to remove these groups because they represent multiple mixed taxa that are likely already strongly represented by other taxonomic groupings. For example, the majority of the Bivalvia indet. observed at St. Paul's are likely the amethyst gem clam *Gemma gemma*, whose shells were too degraded to allow for confident identification to species level. A full list of raw (i.e. unbinned) and binned taxonomic groupings is provided in Table S1.

Table A1 lists the number of taxa identified at each taxonomic level (e.g. family, genus, species etc.) for binned and unbinned data. For each site, we present total number of taxa, total macrofauna abundance, and average abundance of macrofauna per core calculated from binned and unbinned taxonomic groupings (Table A2). We also present a SIMPER analysis, PERMANOVA, and nMDS plot of taxonomic composition for these data (Table A3, Fig. A1). Patterns differed only marginally between these results and the equivalent analyses performed on binned data, which we present in the main text (Table 3, Fig. 3A).

Table A1. Number of taxa identified at each taxonomic level, for binned and unbinned data

Level	Unbinned	Binned
Phylum	1	1
Sub-phylum	1	0
Class	3	0
Order	3	2
Superfamily	2	1
Family	15	10
Subfamily	2	1
Genus	22	19
Species	27	19

Table A2. Total number of taxa, total abundance of macrofauna (N), and average abundance of macrofauna per core (CA) for each site, based on binned and unbinned taxonomic groupings

	Neddie's Harbour		Newman Sound		St. Paul's	
	Unbinned	Binned	Unbinned	Binned	Unbinned	Binned
No. of taxa	36	25	43	32	42	29
N	4232	2285	2701	2632	9426	5131
CA	176	95.2	81.8	79.8	286	156

Table A3. Taxa that contributed most to the dissimilarity between communities (identified by SIMPER analyses performed between each pair of sites) in descending order according to their relative contributions. Mean density (ind. m^{-2}, rounded to the nearest whole number) ±SD is shown for each taxon at each site

Taxon	Neddie's Harbour	Newman Sound	St. Paul's
Gastropoda indet.	5537 ± 3165	116 ± 161	3896 ± 2152
Naididae indet.	325 ± 549	791 ± 964	2209 ± 1454
Paranais litoralis	647 ± 519	236 ± 317	99 ± 243
Pygospio elegans	1368 ± 658	2875 ± 2104	632 ± 372
Bivalvia indet.	198 ± 161	24 ± 49	5320 ± 4193
Gemma gemma	0	0	5707 ± 4849
Manayunkia aestuarina	815 ± 459	0	0
Monocorphium sp.	954 ± 761	11 ± 62	13 ± 37
Mya sp.	682 ± 457	118 ± 98	58 ± 88
Polydora cornuta	121 ± 112	97 ± 90	711 ± 361
Tellinidae indet.	0	511 ± 310	32 ± 64

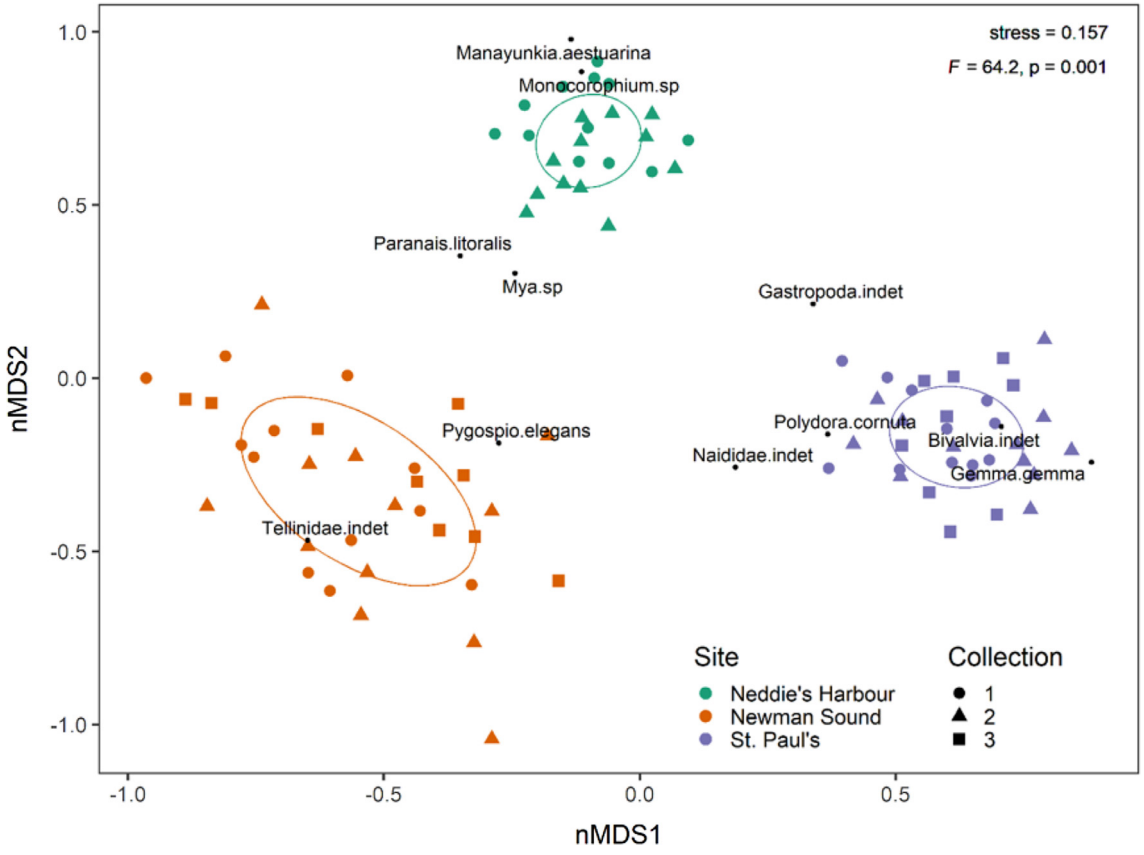

Fig. A1. Non-metric multidimensional scaling (nMDS) plots of taxonomic composition at each site based on unbinned data. Coloured symbols represent individual sediment cores. Colour represents site and shape represents collection ID. Small black dots represent individual taxa identified by SIMPER. Taxa most strongly influence the taxonomic composition of cores to which they are most closely situated. Results of the supporting PERMANOVA (pseudo-*F*, p-value) are displayed in the top right corner of each plot

Editorial responsibility: Erik Kristensen, Odense, Denmark
Reviewed by: P. Renaud, H. Link

Submitted: August 22, 2023
Accepted: March 21, 2024
Proofs received from author(s): April 22, 2024