Detecting strong spatial and temporal variation in macrobenthic composition on an urban shelf using taxonomic surrogates

Broc S. Kokesh[1,*], Susan M. Kidwell[1], Adam Tomašových[2], Shelly M. Walther[3]

[1]Department of Geophysical Sciences, University of Chicago, Chicago, IL 60637, USA
[2]Earth Science Institute, Slovak Academy of Sciences, 84005 Bratislava, Slovakia
[3]Los Angeles County Sanitation Districts, Whittier, CA 90601, USA

ABSTRACT: Surrogates of macrobenthic assemblages, intended to alleviate the effort and taxonomic expertise required for monitoring, can take many forms, such as using coarser taxonomic levels ('sufficiency') or only a subset of the whole fauna ('subsetting'). Here, the power of both approaches to retain community-level patterns of spatial and temporal variation were evaluated using an exceptionally long (47 yr) infaunal dataset generated from monitoring wastewater impacts on an urban shelf in southern California, USA. Four taxonomic sets (whole infauna, polychaetes, bivalves, malacostracans) were evaluated at 5 resolutions (species, genus, family, order, functional guild) along a pollution gradient subdivided into 2 spatial bins based on proximity to the wastewater outfall (near-field vs. far-field) and 3 temporal bins based on wastewater treatment phases. All taxonomic sets detected weakening of the spatial gradient with improved wastewater treatment, i.e. communities became more similar in richness, evenness, and composition through time, and patterns were robust when coarsened to families or guilds. Polychaetes mirrored ('proxied') whole-fauna patterns most accurately, as expected since they constitute most of the individuals and species. However, bivalves outperformed all other sets in detecting ('indicating') the pollution gradient itself, owing to their breadth of feeding strategies. These results strengthen the consistently positive results from taxonomic coarsening emerging from tests elsewhere and the caveats for taxonomic subsetting: clade strengths serve different objectives. Comparable datasets should exist in environmental agency archives elsewhere, promoting the general surrogacy model. For monitoring programs still in their planning stages, regional insights could be acquired via analogous nested analyses of a single survey.

KEY WORDS: Taxonomic sufficiency · Taxonomic subset · Benthic assemblages · Functional groups · Wastewater pollution · Southern California

1. INTRODUCTION

Intense urbanization of coastlines exerts an array of anthropogenic stresses on natural systems through resource harvest, landscape modification such as dredging and armoring, and the input of excess nutrients and other contaminants (Todd et al. 2019). These inputs include agrochemicals and fertilizers, industrial metals, and polycyclic aromatic hydrocarbons via stormwater, aquaculture, vessels, and municipal wastewater, with the latter usually constituting the largest volume of inputs (Islam & Tanaka 2004). The environmental strain imposed by these human stressors is often evaluated by testing for before–after changes and/or far versus near (control–impact) contrasts in the composition and struc-

*Corresponding author: bkokesh@uchicago.edu

ture of benthic communities, which are sampled as part of sediment quality monitoring (e.g. Warwick 1993, Gray & Elliott 2009, Schiff et al. 2016). Such monitoring effort requires a standardized quantitative sampling effort of both organismal abundance and diversity, covers a broad range of taxonomic clades, and identifies organisms to the species level or other fine-scale operational taxonomic units. Long-term programs extending a decade or more are unusual but provide invaluable insights into biological responses over the history of environmental stress (or stress alleviation) in a region (e.g. Borja et al. 2006, 2016, Stein & Cadien 2009, Gogina et al. 2014, Clare et al. 2015, Schwing et al. 2017, Caswell et al. 2018). Such direct information is a kind of 'gold standard' for assessing the trajectory, and thus by implication the present status, of a community or ecosystem.

However, biological monitoring places a high demand on taxonomic expertise and labor, especially in biologically diverse marine systems (Olsgard & Somerfield 2000, Włodarska-Kowalczuk & Kędra 2007). These expectations may potentially inhibit the establishment of new programs, especially in poorly studied regions where assessment of human impacts for public health and biodiversity conservation is urgent. In principle, these demands could be alleviated by using 'surrogates' of species-level data on the whole fauna if the reliability of such surrogates in detecting the human impacts can be assumed (generalizing from studies elsewhere) and further validated. Beyond the concerns of evaluating biodiversity and ecosystem health, the potential biasing effects of using surrogates is also important in historical ecology, which typically has to focus on the subset of taxa that are most likely to have been reported upon (e.g. commercial and sport taxa; Brind'Amour et al. 2014, Di Minin & Moilanen 2014), and in paleoecology (our motivation), which typically has to rely on the subset of taxa with greatest postmortem durability under ordinary conditions of slow burial that characterize most marine environments (e.g. shelled mollusks, calcifying arthropods, and bryozoans; Schopf 1978, Tyler & Kowalewski 2017).

Surrogates fall into 2 categories: (1) using genera or other taxonomic levels coarser than species ('taxonomic sufficiency' sensu Ellis 1985) and (2) using only a subset of the whole fauna, for example focusing on a single class ('taxonomic subsetting'; Table 1). Both reduce the labor and time required to process a sample after sampling, but are rarely evaluated together.

Taxonomic sufficiency has been the subject of innumerable tests in marine settings, mostly finding that coarser taxonomic resolution suffices to detect both anthropogenic (Warwick 1988, Ferraro & Cole 1992, Buss & Vitorino 2010, Dimitriou et al. 2012, Stark et al. 2014) and natural (Włodarska-Kowalczuk & Kędra 2007, Bevilacqua et al. 2009) spatial gradients (for freshwater examples, see Jones 2008 and Mueller et al. 2013, and for terrestrial invertebrates, see Pik et al. 1999 and Souza et al. 2016). Non-Linnaean classifications such as morphospecies and functional guilds have also demonstrated promising results (Bhusal et al. 2014, Brind'Amour et al. 2014).

Tests of taxonomic subsetting have had more diverse objectives (Moreno et al. 2007). In most studies, authors sought to identify a single clade either able to mirror variation in richness or some other attribute of the whole fauna (i.e. a 'proxy' for the whole fauna; e.g. Olsgard et al. 2003, Magierowski & Johnson 2006, Fontaine et al. 2015) or able to efficiently detect a pollution or other gradient (an 'indicator' group), even if it exhibited poor congruence with the whole fauna (Magierowski & Johnson 2006,

Table 1. Common aspects of surrogacy as defined in this study

Term	Definition
Surrogacy	Reducing the effort required to assess biota, via surveys and monitoring
I. Sufficiency	Coarsening the detail (resolution, acuity) of information
A. Numerical sufficiency	Coarsening the measure of abundance (e.g. presence–absence or rank–abundance rather than number of individuals)
B. Taxonomic sufficiency	Coarsening the taxonomic units (e.g. species-level data aggregated into higher Linnaean ranks or into functional guilds)
II. Subsetting	Narrowing the scope or breadth of information
A. Numerical subsetting	Narrowing the number of specimens assessed (e.g. limit to the first 100 individuals processed)
B. Taxonomic subsetting	Narrowing the taxonomic clades assessed (e.g. limit to a selected clade or functional guild)
i. Proxy subset	Subset that best reflects the whole fauna
ii. Indicator subset	Subset that best detects gradient(s) of concern (e.g. indicator taxon)

Heino 2010). Subset surrogates thus fall into 1 of 2 types: 'proxy subsets' (good for mirroring the whole fauna) and 'indicator subsets' (superior at detecting a particular environment or perturbation; Table 1).

How effective are both taxonomic sufficiency and subsetting modes of surrogacy, as well as their inter-action (i.e. subsets at multiple resolutions) at detecting spatial and temporal gradients in anthropogenic disturbances? Here, an exceptionally long-term (~50 yr) and taxonomically consistent species-level macrobenthic database was used to test whether sur-rogates upheld patterns observed using the whole fauna across a well-known anthropogenic pollution gradient over successive phases of wastewater abate-ment. The dataset, which consists of annual infaunal grab samples from 11 stations in 60 m water depth along the Palos Verdes shelf in southern California (USA) since 1970, encompasses about 600 species from 11 phyla. Previous analyses of the full infaunal dataset, including municipal annual reports, have demonstrated improvement in benthic conditions associated with enhanced wastewater treatment (Stein & Cadien 2009, LACSD 2020).

This analysis takes opportunistic advantage of a dataset generated under a government regulatory requirement to conduct long-term monitoring of ben-thic community health and recovery in response to contamination reductions, and thus was not designed as a formal research experiment. In contrast to a standard before–after–control–impact study de-sign (Underwood 1991, 1994 and refer-ences therein), here the before–after component is a temporal analysis over a history of wastewater effluent pollu-tion reduction (Phase 1 with only pri-mary treatment, Phase 2 with partial secondary treatment, and Phase 3 with full secondary treatment; see Fig. 2). Wastewater discharges to the Palos Verdes shelf began in the 1930s and have only been regulated since the early 1970s, and thus benthic response to the onset and rise of pollution is not known directly (but see a sediment-core-based analysis of bivalve shells; Leonard-Pingel et al. 2019). The con-trol–impact component here is a spa-tial comparison of far-field ('refer-ence') and near-field areas on the shelf relative to the pipe opening, with those 2 areas defined by agency scien-tists on the basis of sediment chem-

istry in the 1970s. Formal controls do not exist: no sta-tions on the shelf have been free of pollution for the entire duration of monitoring. The disadvantages of this necessarily less stringent design are, we believe, outweighed by some unusual advantages, namely (1) an exceptionally long time series (47 yr) with (2) scrupulous taxonomic consistency despite turnover in agency personnel (SCAMIT 2018), (3) a strong expert knowledge base for species biology in the region (Coan et al. 2000, Macdonald et al. 2010, SCAMIT 2018), (4) consistent, multiple sampled sites in both near- and far-field areas over the entire duration, (5) consistent sampling methods, and (6) holding water depth constant.

2. MATERIALS AND METHODS

2.1. Study area

The Palos Verdes shelf in southern California has a long history of receiving wastewater contaminants beginning with the 1937 opening of an ocean outfall in 34 m deep water at White Point; 2 Y-shaped out-falls dispersing wastewater at 60 m depth were later installed in 1956 and 1966 (Fig. 1). The 60 m outfalls remain in use, and since 1970, in anticipation of the
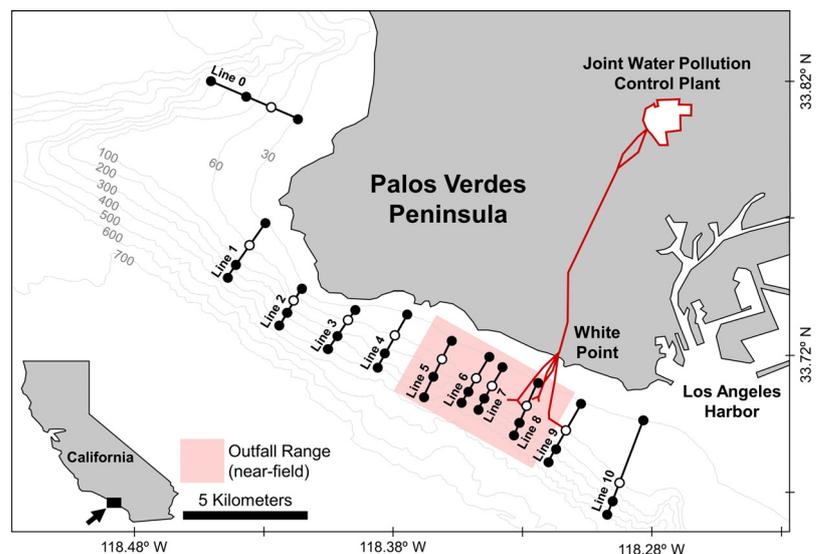
Fig. 1. Study area off the Palos Verdes Peninsula on the southern California (USA) continental shelf. Annual monitoring of macrobenthos is conducted along 11 bathymetric sampling transects (Lines 0–10) at 30, 60, 150, and 300 m depths. White dots indicate sampling stations along the 60 m isobath used in this study. Gray lines are isobaths and red lines are the outfall pipes extending from the Joint Water Pollution Control Plant (JWPCP) to White Point. Pink area approximates the stations near the outfall source (near-field) defined in this study, encompassing Lines 5–8, and the rest of the shelf is considered far-field (Lines 0–4, 9–10). Modified from LACSD (2020)

passage of the US Clean Water Act in 1972, waste-water treatment processes have continually improved. Wastewater was initially subject only to primary treatment, and emission rates steadily increased with 20th century urbanization, with total suspended solids loadings peaking at approximately 150 000 t yr^{-1} (Fig. 2). Emissions of solids (to which most contaminants adhere) began to decrease markedly after the onset of advanced primary treatment in the 1970s, with additional declines in solids and in biological oxygen demand following the onset of partial secondary treatment in the 1980s and 1990s and full secondary treatment starting in the 2000s. Despite a high volume of water outflow to the ocean over the last 50 yr, contaminants in the effluent are now well below allowable emission limits (LACSD 2020).

This emission history created a strong point-sourced spatial gradient in sediment contamination (Swartz et al. 1991, Eganhouse & Pontolillo 2000) that extends mostly northwest of the outfall area, deflected by the California Coastal Countercurrent (Fig. 1). The sediments associated with the effluent discharge form a fine-grained, organic-rich mound as thick as 60 cm near the outfall and contain numerous contaminants including DDT and PCBs (Stull et al. 1996, Lee et al. 2002). This contaminated sediment layer thins

with distance from the outfall; stations where a contaminated sediment layer was detected are designated as near-field stations and those without such a layer are denoted as far-field (not formal controls, because they likely experienced direct or indirect effects from contaminants such as excess nitrogen in the fine sediments released prior to the onset of refined treatment technologies; LACSD 2020). Along the 60 m isobath today, sediment sampled from the immediate near-field area is chemically undistinguishable from that in the far-field stations to the northwest and southeast (Stein & Cadien 2009, LACSD 2020).

Long-term monitoring of macrobenthic communities using benthic grabs (described in Section 2.2) were conducted on the Palos Verdes shelf by the Los Angeles County Sanitation Districts (LACSD) starting in 1970 and are ongoing. Their sampling grid comprises 44 stations: 11 bathymetric transects each sampled at 30, 61, 152, and 305 m water depths (Fig. 1). Infaunal and other data from this standardized effort have revealed a significant and positive benthic response to improved wastewater treatment, including increasing macrofaunal species richness and evenness (LACSD 2020). The faunal composition has, throughout the grid, reached or increasingly
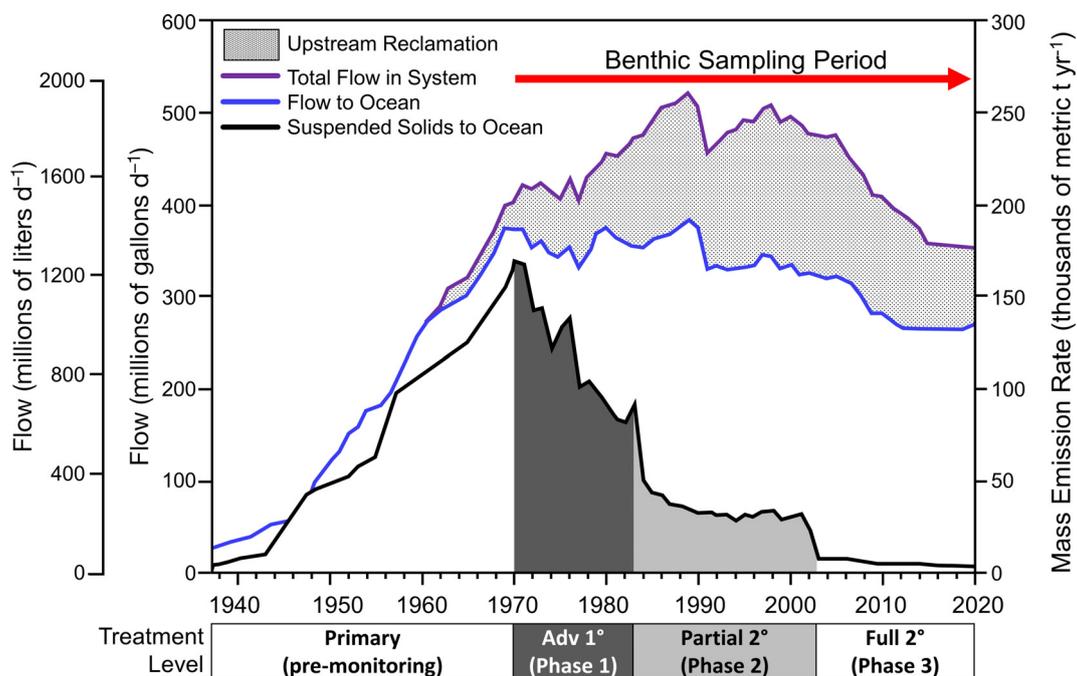
Fig. 2. Discharge history (1 US gallon ≈ 3.79 l) from the Joint Water Pollution Control Plant (JWPCP) through the White Point outfall system from 1937 to 2019. Suspended solid release (10^3 t yr^{-1}) increased until the 1970s and then declined with advanced primary wastewater treatment (Phase 1), partial secondary treatment (Phase 2), and full secondary treatment (Phase 3). Modified from LACSD (2020)

approached a reference condition as defined by the benthic response index, a metric based on the pollution tolerance and abundances of local species calibrated to regional disturbance gradients (Smith et al. 2001), and as compared to other areas within the Southern California Bight that are considered to be undisturbed.

2.2. Dataset

LACSD sampled macrobenthic invertebrate communities from their full spatial grid from 1972 to the present, with the dataset used here encompassing 1972–2019. Samples were collected semi-annually in February and August until 2006, after which the agency switched to annual sampling during the summer. Samples were initially collected using Shipek grabs, which were replaced by van Veen grabs after 1980; here, Shipek grab data have been transformed to equal the volume of van Veen grabs by pooling 3 replicate Shipek grabs taken at each station. Sediments were sieved through 1 mm mesh with seawater on deck, fixed in 10% formalin, and then preserved in ethanol.

Agency biologists sorted and identified all invertebrates to the lowest level possible, i.e. species. Taxonomic identities have consistently been standardized to a regional listing compiled and regularly updated by the Southern California Association of Marine Invertebrate Taxonomists (SCAMIT 2018), a volunteer group of regional professionals; the SCAMIT listing is used by all regional municipal agencies and in the Southern California Bight Regional Monitoring Program (e.g. Schiff et al. 2016). All species assignments in the dataset were updated to a single, internally consistent taxonomy (SCAMIT 2018) for the present analysis.

The analyses for this study focused only on samples from the 60 m isobath, which coincides with the effluent ports of the outfall that have been active point sources since the 1950s. Each of the 11 stations along this middle-shelf isobath was categorized as either 'near-field' (Stns 5–8) or 'far-field' (Stns 0–4, 9–10) based on sediment chemistry conditions during the first decade of monitoring (Stull et al. 1996), an assignment that reflects physical proximity to the outfall and the prevailing northwest-directed California Countercurrent (Fig. 1). Samples collected in February during the years 1972–2005 were excluded, yielding a dataset that consisted of 1 summer-time sample per station per year from 1972 through 2019 for a total of 1939 samples.

This 47 yr history was subdivided into 3 temporal phases based on the treatment level of wastewater emissions: primary and advanced primary (Phase 1 in 1972–1983), partial secondary (Phase 2 in 1984–2002), and full secondary (Phase 3 in 2003–2019) (Fig. 2). At each station, samples were pooled by these intervals to create 3 temporal bins (phases). The final dataset thus consisted of 3 temporal bins at each of 11 stations, which were assigned to one of the 2 outfall categories (spatial bins, referred to as near-field and far-field areas in Section 3). Although seabed conditions clearly improved throughout the Palos Verdes sampling area over these decades of improving wastewater treatment (LACSD 2020), a decision was made to leave stations in their outfall categories as originally defined based on conditions in 1972, allowing us to test the ability of surrogates to detect weakening of the spatial gradient.

Given the stepwise decrease of emissions across the boundaries of treatment phases (Fig. 2), the robustness of datasets to possible lags in benthic responses was tested by omitting the initial 1 or 2 years of data immediately following a change in treatment. Omitting these data did not affect results significantly, and so all years of data were retained in analyses.

2.3. Data analyses

All data extraction and analyses were conducted using R version 4.0.3 (R Core Team 2020) with functions supplied by the 'vegan' package (Oksanen et al. 2020). The original dataset was transformed to 20 surrogate matrices (henceforth called 'sets') for coarsened taxonomic levels and selected taxonomic subsets. First, species-level counts for the whole fauna were aggregated into genus, family, order, and functional guild level counts (tests at 5 levels of taxonomic resolution). Species were assigned to functional guilds (simply 'guilds' hereafter) following Macdonald et al. (2010), who grouped East Pacific species on the basis of animal motility, life habit, food source, diet, and feeding mode. Species in our dataset that were not listed by Macdonald et al. (2010) were omitted from guild analyses. Although guild classifications are often found to be highly conserved within and among related clades (e.g. bivalve families; Stanley 1970, Collins et al. 2019), they should not be viewed as equivalent to Linnean ranks. Using the assignment data from Macdonald et al. (2010), 33 (10%) of the total 335 families in common with the list of Macdonald et al. (2010) contained more than 1

guild (Table 2). However, no single genus in our dataset contained more than 1 guild. Thus, our analysis of guilds represents a coarsening of resolution between the levels of genus and family, but that coarsening is only slightly less than that represented by families.

After taxonomic coarsening the species-level data into these 5 ranked-based matrices, new matrices were extracted from each for the 3 most abundant

Table 2. Representation of taxa from the whole fauna in the functional guild list by Macdonald et al. (2010). Taxa lacking a functional guild assignment were not included in the analysis of guilds. No genera represented more than one guild, although several families did. Guilds thus represent a taxonomic resolution between the levels of genus and family

	Richness	Taxa not assigned a functional guild	Taxa representing a single functional guild	Taxa with >1 functional guild
Species	1277	20	1257	
Genus	722	19	703	0
Family	347	12	302	33
Order	120	11	74	35
Class	34	2	13	19
Phylum	13	0	4	9

classes represented, namely Polychaeta, Bivalvia, and Malacostraca (3 taxonomic subsets, to be compared with the whole-fauna set). These 3 clades together represent ~90% of individuals in the whole fauna (Fig. 3). We chose to assess this many subsets in order to determine whether some behaved as proxy subsets and others as indicator subsets (Table 1). All procedures that follow were applied to each of these 20 matrices: 5 resolutions of information for 3 taxonomic subsets and for the whole fauna.

The distribution of the number of species within the 5 supra-specific categories (i.e. groupings higher than species level) was quantified as the ratio between the raw number of species (i.e. richness, S) over the number of genera (G), families (F), orders (O), and functional guilds (FG) (i.e. S/G, S/F, S/O, S/FG). The smallest possible value of any ratio is 1, which indicates that the higher taxon is monospecific.

Richness and evenness were compared between near-field and far-field samples for each treatment phase by comparing the median values and interquartile range (IQR) of one area to the other. Richness was rarefied to the smallest sample size per taxon per treatment phase to account for unequal abundances among the taxonomic sets (3727 individuals for the whole fauna, 1783 Polychaeta, 285 Bivalvia, 44 Malacostraca). Evenness is expressed as the probability of interspecific encounter (PIE; Hurlbert 1971), which has low sensitivity to sample size and richness. PIE values have potential to range from 0 to 1, with 0 indicating a sample where all individuals are from a single taxon (i.e. the chance of 2 randomly selected individuals being different taxa is zero), and with a PIE value of 1 indicating that individuals are equally distributed among all taxa.

Compositional variation among samples was assessed using non-metric multidimensional scaling (NMDS) based on Bray-Curtis distances computed with square-root transformed proportional abundances (Hellinger transformation). Bray-Curtis values

Fig. 3. Proportional (A) abundances and (B) species richness among the 10 most abundant faunal classes represented in benthic samples from 1972 to 2019. Numbers within bars are the raw numbers of individuals or species. The total abundance (N) and species richness (S) for the whole fauna are printed in the upper right corners. White bars in (B) indicate 3 classes that rank among the 10 most abundant (in A) but are not among the top 10 in richness. Our selected taxonomic subsets (Polychaeta, Bivalvia, and Malacostraca) represent 88% of sampled individuals and 72% of sampled species during the 47 yr of monitoring along the 60 m isobath of the Palos Verdes shelf

quantify the compositional dissimilarity among stations on a scale from 0 to 1 (Bray & Curtis 1957). NMDS ranks these distances prior to ordination on a fixed number of axes (here, $k = 2$). Compositional differences among near-field and far-field stations for each treatment phase were also tested using permutational multivariate analysis of variance (PERMANOVA; Anderson 2001). R^2 values, generated by PERMANOVA, which measure the amount of variation explained by sample categorization of near-field and far-field conditions, were compared among sets to assess the relative strengths of surrogates to delineate the near-field region of the shelf during each phase of wastewater treatment.

The contributions of individual taxa to the Bray-Curtis dissimilarity between near-field and far-field communities for each set were calculated using the similarity percentage procedure (SIMPER; Clarke 1993). To quantify the proportion of taxa that contribute most to dissimilarity for each set, the number of taxa that met increasing levels of cumulative contribution (25, 50, and 75%) was divided by the total number of unique taxa. Higher proportions at contribution thresholds indicate that a higher number of taxa is needed to account for the observed dissimilarity.

3. RESULTS

3.1. Taxonomic structure of the macrobenthic fauna

The 'whole fauna' dataset comprised 536 056 individual organisms across 1277 species, 722 genera, 347 families, 120 orders, 34 classes, and 13 phyla, and these species encompass 86 guilds (Table 3). The 3 focal classes comprised 89% of all individuals, with polychaetes encompassing 61%, bivalves 17%, and malacostracans 10% (Fig. 3A). These 3 classes also comprised 72% of all species: 42% for polychaetes, 7% for bivalves, and 23% for malacostracans (Fig. 3B). Seven of the 10 most abundant classes were also the highest in species richness.

Polychaetes were the most speciose per higher taxon and bivalves were the least speciose subset (Table 3). For example, the S/G ratio for polychaetes was 2.2 as opposed to 1.8 for the whole fauna, and 1.4 for bivalves. Per-family species richness for polychaetes (S/F = 11) was ~3 times higher than other subsets and the whole fauna (S/F = 3.1–3.7), and per-order species richness of both polychaetes and malacostracans was ~3 times higher than the whole fauna (S/O = 33.7, 36, and 10.6, respectively). Bivalves were the least speciose subset at all levels, with the contrast increasing at successively higher levels (Table 3).

Considering the whole fauna, guilds included larger numbers of species than did orders (S/FG = 14.9, S/O = 10.6; Table 3). In each taxonomic subset, the S/FG ratios typically fell between the S/F and S/O levels. The polychaete S/FG ratio (15.8) most closely resembled that of the whole fauna (14.9), while those of bivalves and malacostracans were notably smaller (S/FG of 7.4 and 9, respectively).

3.2. Temporal and spatial variation in richness

Temporally, rarefied species richness increased with improved wastewater treatment in all 4 taxonomic sets — the whole fauna and 3 class-level subsets (first column in Fig. 4). Richness of the whole fauna, polychaetes, and bivalves exhibited stepwise increases across each phase boundary; a gain of ~100 species by the whole fauna, ~50 by polychaetes, and ~10 by bivalves. Malacostracans exhibited a large increase from Phase 1 (1972–1983) to Phase 2 (1984–2002), but negligible change from Phase 2 to Phase 3 (2003–2019).

Table 3. Taxonomic richness at 7 hierarchical levels and the ratio between the number of species and number of supra-specific taxa (S/T) for the whole benthic fauna and 3 taxonomic subsets. Higher ratios indicate higher within-taxon diversity. Empty cells have a value of 1

Hierarchical level	Whole		Polychaeta		Bivalvia		Malacostraca		Average
	S	S/T	S	S/T	S	S/T	S	S/T	S/T
Species (S)	1277		539		89		288		
Genus (G)	722	1.8	247	2.2	62	1.4	166	1.7	1.8
Family (F)	347	3.7	49	11	29	3.1	89	3.2	5.3
Order (O)	120	10.6	16	33.7	10	8.9	8	36	22.3
Class (C)	34	37.6							
Phylum (P)	13	98.2							
Functional Guild (FG)	86	14.9	34	15.8	12	7.4	32	9	11.8
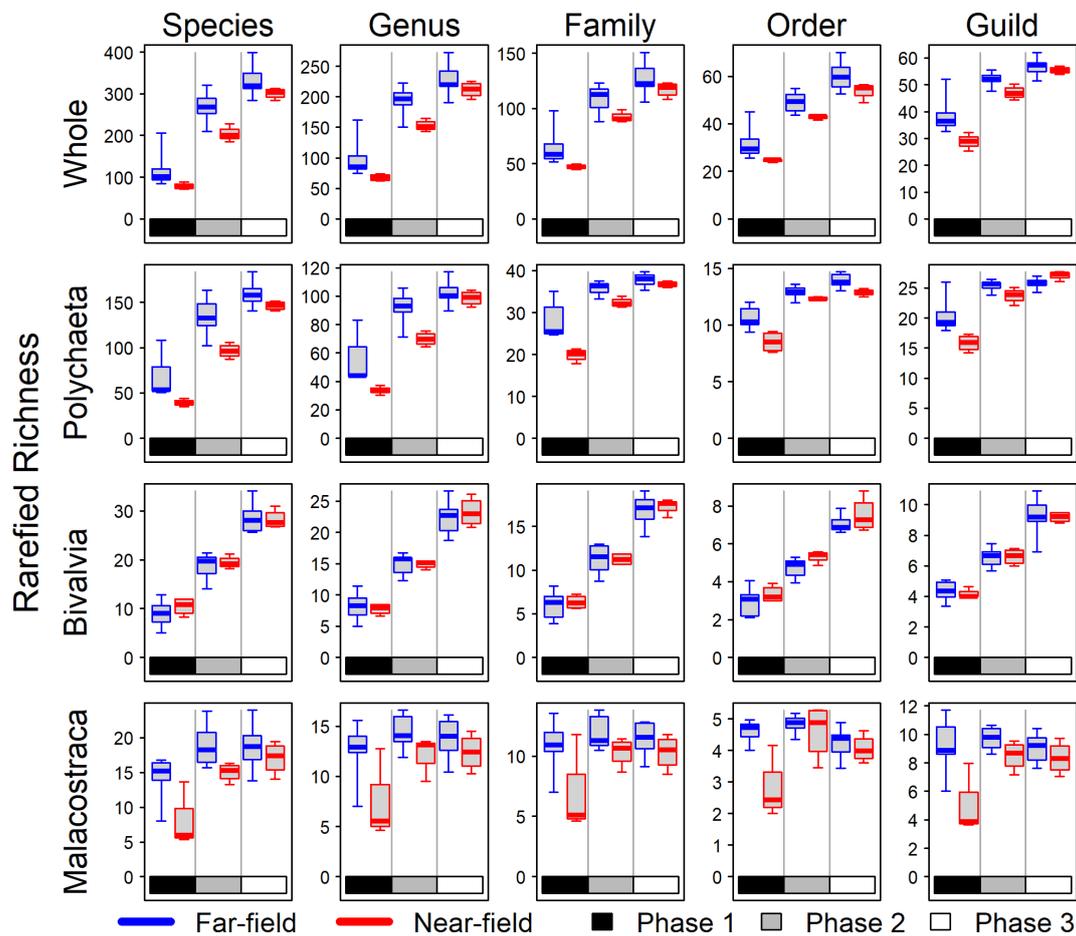
Fig. 4. Rarefied richness for whole benthic samples and subsets (rows) and coarsened taxonomic resolutions (columns). Heavy line is median, box denotes the interquartile range (IQR), and whiskers denote entire range of values. Paired boxes compare stations at far-field (blue; 7 stations at 60 m depth) and near-field (red; 4 stations) areas on the shelf, as judged from sediment chemistry during the 1970s. In each graph, richness was rarefied to the smallest station-level sample size. Richness increases with improved wastewater treatment in all analyses, from Phase 1 (1974–1983) to Phase 2 (1984–2002) and Phase 3 (2003–present; shaded bars along *x*-axes). Polychaete patterns best mirror the whole fauna, and bivalves exhibit the strongest increase across phase boundaries. For all sets except bivalves, richness is higher at far-field stations rather than near-field. This contrast is damped at coarser taxonomic levels

Spatially, with the exception of the bivalve subset, species richness was lower at near-field than at far-field stations, with the contrast strongest during Phases 1 and 2 and smallest during Phase 3 (Fig. 4). Bivalves exhibited no difference in richness between near-field and far-field stations during any of the treatment phases.

These species-level differences in richness over time and between areas persisted with taxonomic coarsening to the family level, although family-level polychaete data dampened the magnitude of temporal increase across each phase boundary. Coarsening data to the ordinal level preserved patterns for the whole fauna, but significantly altered patterns for each subset, particularly for malacostracans (Fig. 4). Guilds appear to have the opposite effect from orders: using guild-level data altered patterns for the

whole fauna, but preserved patterns, albeit slightly dampened, for each of the subsets (Fig. 4).

3.3. Temporal and spatial variation in evenness

With improved wastewater treatment, species-level evenness generally increased and variation among stations within both the near-field and far-field areas generally decreased (i.e. smaller IQRs); these changes were observed in all sets except malacostracans, which maintained consistently high evenness (Fig. 5). Bivalves exhibited the strongest increase in evenness over time — stronger than the whole fauna — and polychaetes most closely resembled the whole fauna, but all sets reached an evenness > 0.9 by Phase 3.
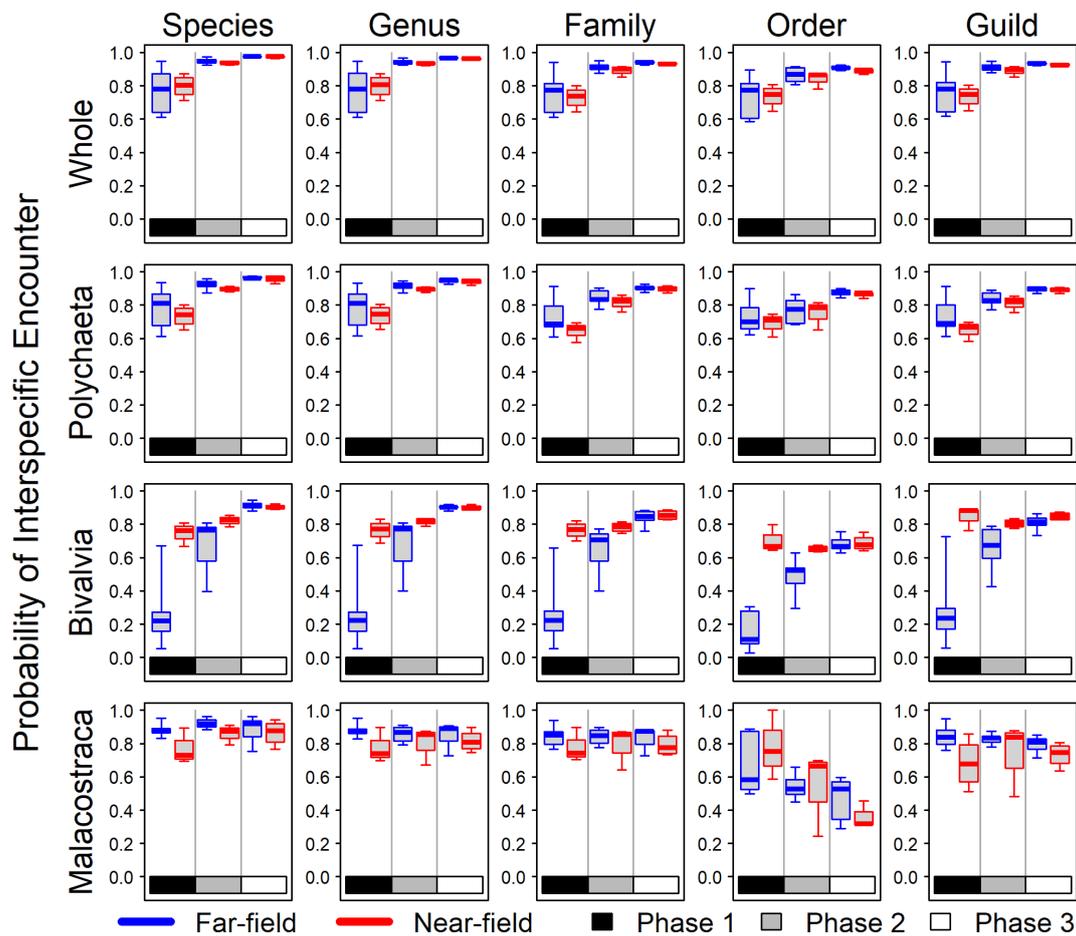
Fig. 5. Taxonomic evenness (probability of interspecific encounter) for whole benthic samples and subsets (rows) at increasingly coarse taxonomic resolutions (columns), with plots organized as in Fig. 4. Evenness increases with improved wastewater treatment, but bivalves exhibit an especially strong trend arising largely from strong contrasts between near-field and far-field stations early in the treatment history. Patterns persist with taxonomic coarsening (left to right within each row) except for malacostracans at the ordinal level

Spatially, during each treatment phase, polychaetes mirrored patterns exhibited by the whole fauna with no significant difference in evenness between near-field and far-field stations. In contrast, bivalves had far higher evenness at near-field than far-field stations during both Phase 1 and Phase 2 (Fig. 5). By Phase 3, these differences among bivalves disappeared, with evenness values from both near-field and far-field stations reaching 0.9–1.0.

As with richness, species-level patterns persisted with coarsening of taxonomic resolution up to the family level (left to right across Fig. 5). Ordinal-level data did not change patterns in the whole fauna but did alter polychaete and bivalve patterns (mostly a decline in maximum evenness). Malacostracans had greater ranges of evenness and a particularly strong temporal decline at the ordinal level. Guilds preserved similar patterns as the family level for the

whole fauna, polychaetes, and bivalves, but slightly altered malacostracan patterns.

3.4. Temporal variation in strength of spatial gradient

At the species level, ordination revealed a consistent pattern of temporal and spatial separation among samples (Fig. 6; each icon is a station within either the near-field or far-field area during a treatment phase). Generally, successive water treatment phases separated along NMDS axis 1, while near-field and far-field stations within each phase separated along NMDS axis 2. All ordinations revealed a decrease in multivariate dispersion over time, namely smaller convex hulls, i.e. less compositional variation among sampled stations within an area, as well as progres-
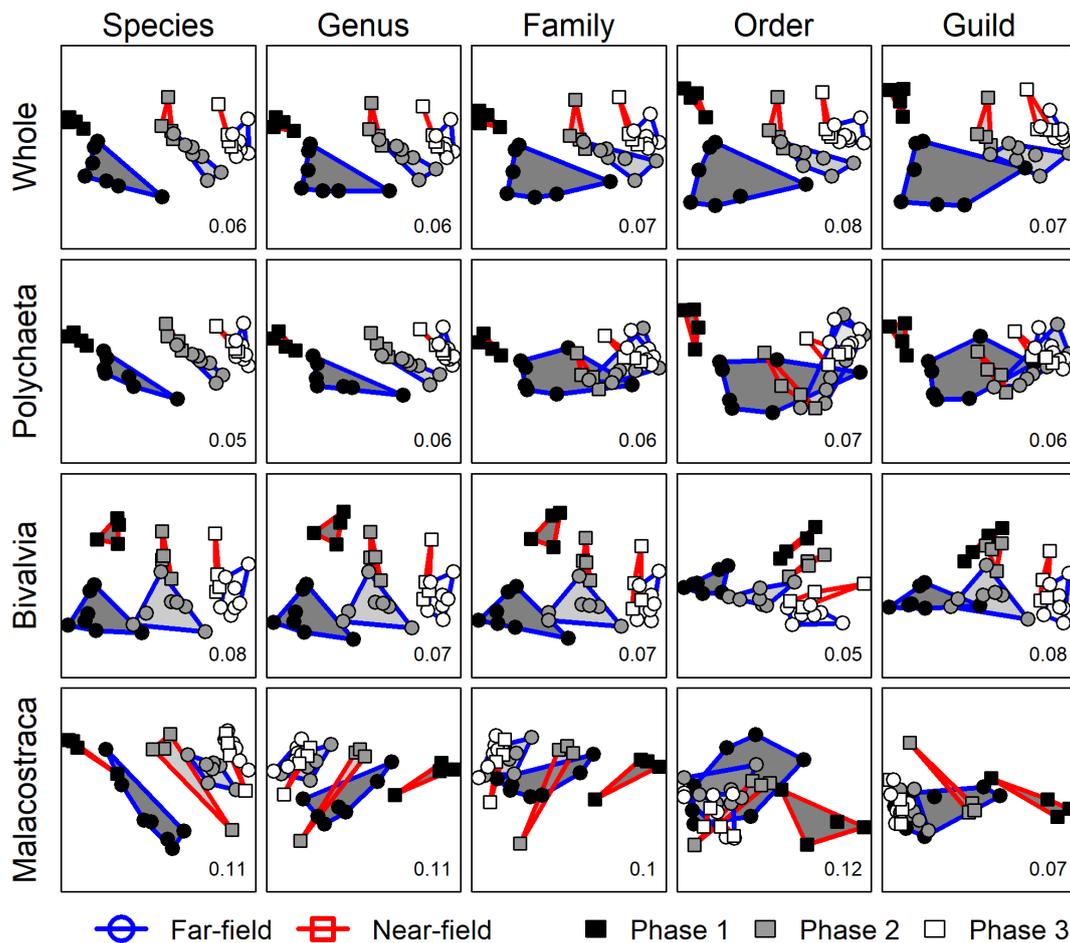
Fig. 6. Non-metric multidimensional scaling (NMDS) plots of Bray-Curtis similarities for whole benthic samples and subsets (rows) at increasingly coarse taxonomic resolution (columns). Distances are based on square-root proportional abundances; convex hulls group samples by water treatment phase (point colors) and outfall proximity (point symbols); ordination stress values printed in lower right corner of each plot. All analyses approximately ordinate successive treatment phases along NMDS axis 1 (horizontal axis) and ordinate the spatial gradient along NMDS axis 2 (vertical axis). The taxonomic compositions of near-field and far-field stations become increasingly similar as wastewater treatment improves, both in the whole fauna (top row) and in each subset

sive convergence in community composition of the near- and far-field areas (a decrease over time in beta diversity along the 60 m isobath).

These species-level patterns persisted for all sets up to the family level and also for guilds. Coarsening resolution to the ordinal level changed the relative positions of samples, but temporal translation and spatial homogenization are still apparent.

These temporal changes were also detected by PERMANOVA. Near- and far-field areas differed significantly in composition during Phase 1 and less strongly over successive treatment phases, becoming insignificant during Phase 3 (Fig. 7). Together with NMDS (Fig. 6), PERMANOVA results indicated spatial homogenization in community composition and a loss in strength of the spatial pollution gradient over time. R^2 values for polychaetes and mala-

costracans were similar to the whole fauna, progressively declining from ~0.45 during Phase 1 to 0.3 during Phase 2 and <0.2 during Phase 3. Bivalves demonstrated both the strongest early differences ($R^2 = 0.65$ during Phase 1) and the steepest decline across treatment phases (0.4 during Phase 2, <0.1 during Phase 3). Resolution had little impact on PERMANOVA results up to the family level and for guilds: at the ordinal level, the whole fauna pattern was unaffected, significant differences were detected for polychaetes during Phase 3, differences were even stronger for bivalves during Phase 1 and Phase 2, and differences for malacostracans during Phase 2 were no longer detected. Ordinal-level data thus significantly affected PERMANOVA results.

SIMPER identified that 2 species are most responsible for spatial variation in the composition of the

Fig. 7. PERMANOVA R^2 values for compositional differences between near-field and far-field stations during each phase of wastewater treatment (bar color) for whole benthic samples and subsets (rows) at increasingly coarse taxonomic resolution (*x*-axis). Asterisks indicate significant p-values (*p < 0.05, **p < 0.01). All taxonomic sets at all taxonomic resolutions detect spatial homogenization of community compositions with improved wastewater treatment. Polychaetes were most consistent with the whole fauna and bivalves exhibited the strongest change in R^2 values across treatment phases
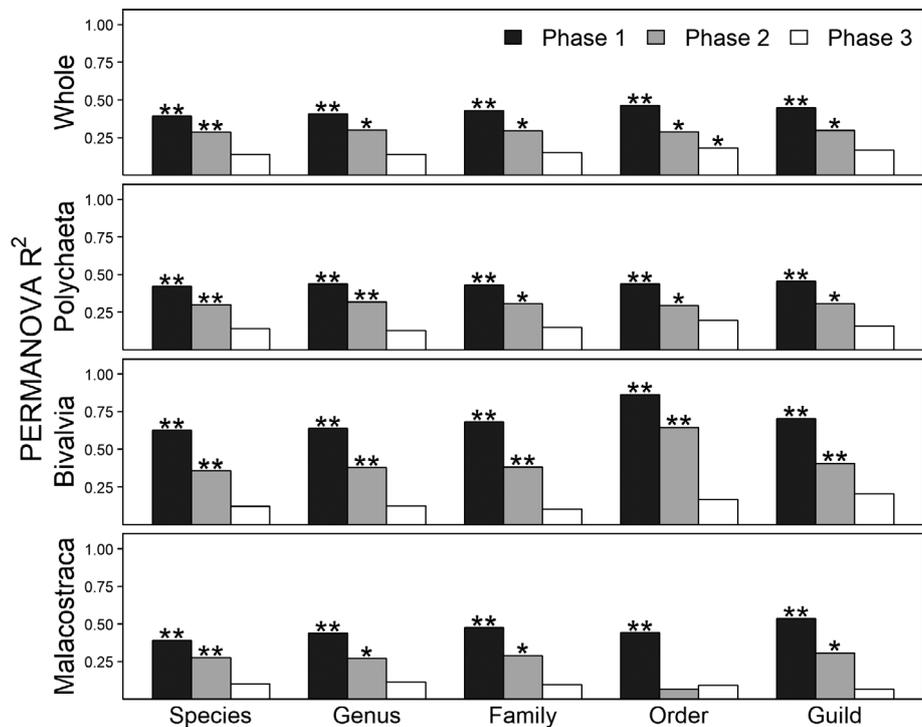
whole fauna during Phase 1, namely the opportunistic polychaete *Capitella capitata* (9% contribution) and the chemosymbiont-bearing bivalve *Parvilucina tenuisculpta* (7% contribution). All other species had <4% contribution. Abundances of *C. capitata* were much higher in the near-field area during Phases 1 and 2 while *P. tenuisculpta* was more abundant in the far-field area (Fig. 8). Both species declined exponentially in numerical abundance over successive wastewater treatment phases (from densities >10^3 to <10^2 for both species). By Phase 3, neither species was a dominant contributor to spatial dissimilarity, which had also become insignificant for the whole fauna (white bars in Fig. 7).

Richness increased from Phase 1 to Phase 2 for all sets at all taxonomic resolutions but remained mostly stable from Phase 2 to Phase 3 (Total S, Table 4). However, SIMPER for the whole fauna, polychaetes, and bivalves indicated that spatial homogenization of communities, i.e. the weakening of spatial gradients, continued from Phase 2 to Phase 3 (parenthetical values; Table 4), with decreasing per-taxon contributions to dissimilarity across all 3 treatment phases. These patterns persisted with coarsening taxonomic resolution up to the ordinal level. Patterns for malacostracans were less consistent: while the number of sub-ordinal taxa increased during each phase, the proportion of taxa that notably contributed to dissimilarity decreased.
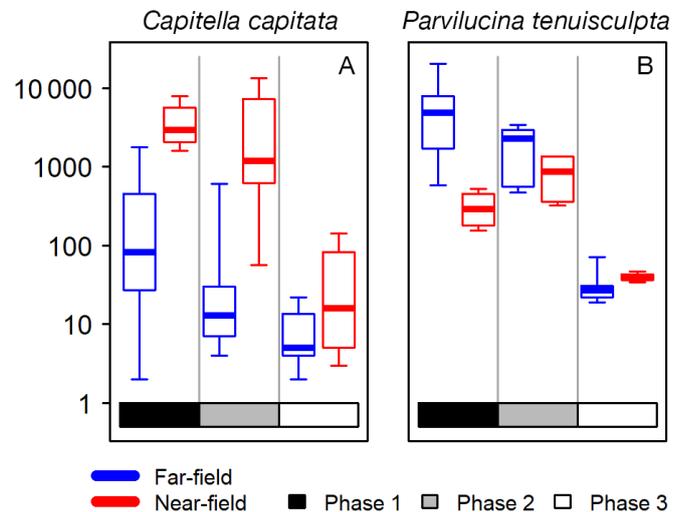
Fig. 8. Raw abundances of the (A) polychaete *Capitella capitata* and (B) bivalve *Parvilucina tenuisculpta*. Paired boxes compare near-field and far-field stations as described in Fig. 4. *C. capitata* abundance is higher at near-field stations while *P. tenuisculpta* is more abundant at far-field stations. As wastewater treatment improved, the abundance of both species decreased and became more similar spatially

4. DISCUSSION

A vast number of analyses on taxonomic surrogacy have been conducted in the past, spanning numerous locations, environmental conditions, and faunas

Table 4. Summary of SIMPER results for each dataset over 3 treatment phases, indicating the number of taxa that cumulatively contribute to 25, 50, and 75% of total Bray-Curtis dissimilarity. Parenthetical values are the proportion of the total number of taxa (S) in a dataset. Higher proportions indicate that more taxa are required to explain the same amount of dissimilarity among samples, thus dissimilarity is lower

Cumulative contribution	Species			Genus			Family			Order		
	Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3
Whole fauna												
25%	6 (0.02)	28 (0.03)	54 (0.06)	4 (0.01)	17 (0.03)	30 (0.04)	3 (0.02)	9 (0.03)	15 (0.05)	2 (0.02)	4 (0.04)	7 (0.07)
50%	28 (0.08)	111 (0.11)	169 (0.18)	21 (0.08)	64 (0.11)	98 (0.14)	12 (0.08)	31 (0.11)	46 (0.16)	5 (0.04)	10 (0.10)	18 (0.17)
75%	91 (0.24)	288 (0.30)	363 (0.38)	67 (0.24)	163 (0.28)	214 (0.30)	37 (0.24)	78 (0.28)	103 (0.36)	14 (0.12)	26 (0.26)	39 (0.37)
Total S	373	975	953	279	579	723	155	282	285	121	99	106
Polychaeta												
25%	4 (0.02)	13 (0.02)	27 (0.06)	3 (0.02)	7 (0.03)	14 (0.06)	3 (0.06)	3 (0.06)	5 (0.11)	2 (0.12)	2 (0.12)	3 (0.20)
50%	16 (0.09)	53 (0.10)	82 (0.19)	12 (0.10)	27 (0.13)	43 (0.20)	7 (0.14)	9 (0.19)	13 (0.28)	3 (0.18)	4 (0.24)	5 (0.33)
75%	50 (0.27)	132 (0.24)	173 (0.39)	36 (0.29)	65 (0.31)	89 (0.40)	15 (0.30)	19 (0.40)	24 (0.51)	6 (0.35)	7 (0.41)	9 (0.60)
Total S	183	540	440	125	221	220	50	47	47	17	17	15
Bivalvia												
25%	2 (0.05)	4 (0.05)	6 (0.10)	1 (0.03)	3 (0.06)	5 (0.08)	1 (0.05)	2 (0.07)	3 (0.12)	1 (0.12)	1 (0.11)	2 (0.20)
50%	4 (0.10)	9 (0.12)	15 (0.25)	3 (0.09)	7 (0.13)	12 (0.19)	3 (0.16)	4 (0.13)	8 (0.31)	2 (0.25)	2 (0.22)	3 (0.30)
75%	8 (0.20)	21 (0.29)	29 (0.48)	6 (0.019)	14 (0.27)	23 (0.37)	4 (0.21)	8 (0.27)	14 (0.54)	2 (0.25)	3 (0.33)	6 (0.60)
Total S	40	73	61	32	52	63	19	30	26	8	9	10
Malacostraca												
25%	4 (0.06)	8 (0.04)	13 (0.04)	3 (0.05)	4 (0.03)	6 (0.04)	3 (0.07)	3 (0.04)	4 (0.04)	2 (0.29)	1 (0.12)	1 (0.12)
50%	9 (0.13)	25 (0.12)	35 (0.12)	7 (0.12)	15 (0.12)	19 (0.11)	6 (0.14)	9 (0.12)	12 (0.13)	3 (0.43)	3 (0.38)	3 (0.38)
75%	23 (0.32)	65 (0.31)	71 (0.25)	15 (0.27)	38 (0.29)	41 (0.25)	11 (0.26)	22 (0.30)	27 (0.30)	4 (0.57)	5 (0.62)	4 (0.50)
Total S	71	213	289	56	130	167	43	74	90	7	8	8

(Moreno et al. 2007, Bacci et al. 2009). This wealth of information and some meta-analyses (Mellin et al. 2011, Westgate et al. 2014) indicate that the usefulness of surrogates depends on the study system (e.g. substrate, latitude, strength of the environmental gradient, spatial scale), the study aim (e.g. detection of gradients, estimation of diversity, richness hotspots), and the methodologic focus (taxonomic sufficiency, which tends to yield positive results from coarsening resolution, versus taxonomic subsetting, which yields highly varied results).

The Palos Verdes shelf evaluated here is a single study system where surrogates might be unusually effective. However, our multi-decadal time-series of annual sampling with exceptional taxonomic consistency permits an exceptional breadth of tests without the challenges of comparing among published studies. Comparable datasets might well exist in the archives of environmental management agencies elsewhere, permitting counterpart analyses to support a general model of surrogacy (e.g. shelves with differing productivity, tropical shelves, coastal embayments). In addition to spurring the discovery and academic re-use of such datasets, nested analyses similar to those here could be applied to a pilot survey, allowing managers to evaluate the viability of potential surrogates to satisfy the objectives of a new monitoring program.

4.1. Caveats

The dataset evaluated is itself only a subset of the total macrobenthic fauna in the study area owing to the use of Shipek and van Veen grab samplers, which are most effective at collecting infaunal taxa living within the upper ~15 cm of the seabed (minimum penetration required to retain a sample; LACSD 2020). Although these data largely exclude mobile epifauna, and certainly planktonic meiofauna, the taxa collected by these and other point-pene-

tration sampling gear (e.g. 0.01 m^2 box cores) are the basis of most quantitative analyses of soft-sedimentary communities by both academics and agencies (Eleftheriou & Moore 2013). It is unknown how results may vary if we instead or additionally employed sampling methods (trawling, plankton nets, etc.) that better target these other commonly studied aspects of the overall community.

Numerical sufficiency (Table 1), although not assessed here, can significantly reduce the effort of counting individuals in samples (e.g. Carneiro et al. 2010, Landeiro et al. 2012) but limits the analytical utility of data. Presence–absence, ranked abundance, and categorical (abundant, common, rare, etc.) data, for example, will quantify raw richness accurately, but cannot be rarefied and are unsuitable for evaluating evenness (community structure). Such low-resolution data are also problematic for SIMPER analysis, which specifies the use of abundance-based Bray-Curtis dissimilarities (Clarke 1993). For these reasons, the absolute counts of individuals for each sample, i.e. density data, were used in all analyses.

Data transformation (i.e. analytical coarsening) of numerical abundance data constitutes a test of numerical sufficiency, but of course does not reduce processing effort and thus is not a surrogate. Up to the family level, our patterns using Hellinger-transformed data are robust to other transformations of proportional abundances, including no transformation or fourth-root transformation. The main effect of data transformation is on the strength of the spatial difference detected by taxonomic subsets during Phase 1. Although differences are significant during Phase 1 for all surrogates, PERMANOVA R^2 values decline as the transformation intensity increases; the same trend was observed in the other temporal bins. Coarsening numerical resolution to the extreme of presence–absence data makes the patterns robust only to the genus level.

We use the 3 most numerically abundant macrofaunal classes as subsets in this study, but it should be noted that the class Gastropoda ranks slightly above Bivalvia in species richness in this dataset (Fig. 3B). However, gastropods have such low abundance on the Palos Verdes shelf that they did not, on their own, show any spatial or temporal variation. Thus, in serving as surrogates, taxonomic subsets represented by large numbers of individuals outperform subsets with large numbers of species. The analyses in this study thus focused on the 3 most numerically abundant, rather than most speciose, classes as taxonomic subsets.

4.2. Taxonomic resolution: genus- and family-level data suffice

Coarsening taxonomic resolution to genera and families preserved both the spatial and temporal gradients in wastewater contamination that were detected by the whole macrobenthic fauna using species-level data, and the same was true for all assessed surrogate subsets, whereas ordinal-level data only detected a spatial gradient when one existed (i.e. during Phases 1 and 2; Fig. 7). These results corroborate previous studies that taxonomic coarsening to family-level identification is a good first-order means of reducing the effort of detecting spatial gradients (e.g. Jones 2008, Heino 2010, Mellin et al. 2011, Pitacco et al. 2019). Our analysis explicitly extends this result to temporal variation and, more importantly, reveals its (mostly stable) interaction with different taxonomic subsets of the fauna (see Sections 4.3 and 4.4).

Genus- and family-level data probably sufficed in this southern California setting because S/G and S/F ratios in the whole fauna are relatively low (~2 and 5, respectively; Table 3), consistent with studies in other systems (e.g. Bevilacqua et al. 2012, de Oliveira et al. 2020). Sufficiency is commonly found to be inversely related to the ratio of the number of species over the number of higher taxa (S/T): increasing the number of species per higher taxon increases the probability that these species will differ in their pollution tolerance. Some studies postulate that S/T values <2.5 will generally result in good correlations with species-level patterns (Timms et al. 2013, Albano et al. 2016). Although the family and guild ratios in this study are almost all >2.5 (Table 3), the S/T values on the Palos Verdes shelf corroborate the growing evidence that the sufficiency of higher taxa depends on S/T ratios remaining below a particular threshold.

The unpredictable patterns produced by ordinal data across all faunal datasets probably arise at least in part from the high S/O values (8.9–36, Table 3), although the small number of orders probably also contributes to their low sensitivity. In malacostracans, for example, where diversity and compositional patterns deteriorate most strongly with coarsening resolution, the S/T ratio rises strongly from families (S/F = 3.2) to orders (S/O = 36).

The exceptionally strong anthropogenic environmental gradient along the Palos Verdes shelf during Phases 1 and 2 also favors the practicality of coarser groupings: habitats as distinct as black H$_2$S-rich mud and greenish sandy mud are all soft but are likely to be dominated by different guilds rather than simply

by different species within a single, consistently dominant guild. Guilds have been shown to be conserved evolutionarily within families and genera (for bivalves, see Stanley 1970, Collins et al. 2019). Consequently, coarsening taxonomic resolution to guild-level information should generally not reduce the ability to differentiate spatial and temporal variation, especially in areas with known or suspected strong human stressors.

Finally, species-level taxonomy undergoes revision, even in areas with well-resolved macrobenthic systematics such as southern California, challenging the merging of species-level datasets across decades of monitoring. The recognition of new species by splitting older ones creates the greatest complications for historic compilations in the absence of voucher material retained for future reference. For example, in biological monitoring of California shelves, the bivalves *Tellina carpenteri* and '*Tellina* sp. B' were not differentiated prior to the 1990s. Such newly split taxa will increase the S/G ratio, but they constitute a very small minority of taxa in the fauna. Other taxonomic revisions are, in contrast, easily applied retrospectively to data and will tend to decrease the S/G ratio, such as the correction of synonyms and the upgrading of subgenera to genera. The greater challenge is the per-genus species richness hidden by known but unresolved taxonomic complexes, that is, taxa that harbor cryptic species whose distinction is beyond conventional morphologic detection (e.g. the polychaetes *Capitella capitata* and *Scoloplos armiger*). Fortunately, such complexes constitute only a small number of taxa in the temperate southern California fauna. A potential solution to alleviating such challenges, especially in comparisons among different regions, would be the implementation of environmental DNA metabarcoding to identify operational taxonomic units (OTUs) without the biases of phenotypic taxonomic identification (Ruppert et al. 2019). The ecological anonymity of OTUs will, however, represent a significant loss of insight.

4.3. Polychaetes are the best proxy of spatial and temporal variation in the whole fauna

Polychaete patterns resembled the whole fauna most closely, indicating that they are the best proxy subset (Table 1) for characterizing spatial and temporal variation of benthos on the Palos Verdes shelf. This fidelity likely arises from polychaetes constituting the majority of individuals in the whole fauna (61%; Fig. 3, Table 3), not simply a plurality, and per-

sisted with coarsening to the genus and family levels (Figs. 4–7). In contrast, constituting a high proportional richness of the fauna did not make a clade here an inherently effective proxy (e.g. malacostracans and the species-rich but individual-poor gastropods). Thus, although polychaetes are the most speciose class in our fauna, their proxy power results from their high proportional numerical abundance.

For similar reasons, other studies have also found polychaetes to be excellent proxies of the whole fauna (e.g. Giangrande et al. 2005, Magierowski & Johnson 2006, Włodarska-Kowalczuk & Kędra 2007). Subsets finer than a class, however, are highly variable in power. For example, Olsgard et al. (2003) identified the order Terebellida as an effective proxy subset for all polychaetes and, to a lesser extent, the whole fauna in a spectrum of temperate to subarctic shelves. In our California 60 m isobath dataset, Terebellida was the most numerically abundant order (34% of polychaetes and 21% of the whole fauna), but failed to mirror spatial and temporal patterns of either the full polychaete subset or the whole fauna; this family also failed to detect the spatial pollution gradient. Capitellidae, the second most abundant polychaete clade (23% of polychaetes, 14% of whole fauna), also failed to mirror the full polychaete subset or the whole fauna, although it did rival Bivalvia in its ability to detect the spatial gradient (see Section 4.4). The entire class of Polychaeta is therefore an effective proxy subset of the whole fauna in southern California, but subsets within Polychaeta do not appear to be suitable on their own.

4.4. Bivalves are the best indicators of spatial and temporal gradients

Bivalves did not mirror the whole fauna and thus performed poorly as proxies in this study area, but out-performed both the whole fauna and other subsets in detecting the outfall gradient and its temporal decline. Bivalve richness was sensitive to decadal-scale temporal variation, but surprisingly, not sensitive to spatial variation (Fig. 4): the temporal contrast in conditions (between Phases 3 and 1) was greater than the spatial contrast (near-field versus far-field) even within Phase 1, implying that present conditions for bivalves everywhere on Palos Verdes shelf permit more species than even the far-field area during peak pollution. In contrast, bivalve evenness, and composition, i.e. measures that consider numerical abundance, excelled at detecting both spatial and temporal variation on this urban shelf (Figs. 5 & 6).

Lucinida (including *Parvilucina tenuisculpta* and other facultative chemosymbiont-bearing bivalves) was the most numerically abundant bivalve order in this dataset (77% of bivalves, 13% of whole fauna), followed by Venerida (20% of bivalves, 3% of whole fauna; part of the infaunal suspension feeding guild). However, neither clade alone could mirror patterns of the full bivalve subset or strongly detect the spatial gradient. The entire class of Bivalvia is therefore an effective indicator subset for this outfall gradient and, like the polychaetes discussed above, subsets within Bivalvia were not suitable on their own.

4.5. The predictive power of functional guilds for surrogacy

Changes in the relative abundance of trophic guilds have been long-appreciated as a key component of benthic response to pollution, as observed over time, along spatial gradients, and in experiments (e.g. Pearson & Rosenberg 1978, Gray & Elliott 2009, Riedel et al. 2012). Species that feed on plankton and other suspended material (filter feeders and photo-symbiont-bearing taxa) decline in favor of those that feed on detritus and/or its microbial communities (many deposit feeders plus chemosymbiont-bearing taxa; Simonini et al. 2004, Kotta et al. 2007, Villnäs et al. 2012). Many detritus-feeders are additionally tolerant of the side effects of organic enrichment, such as hypoxia, soupy textures, resuspended fine sediments, and contaminants absorbed on fine particles (Maurer et al. 1999). The relative abundance of guilds that consider mobility, life habit, and trophic group is thus expected to vary significantly along environmental gradients, and perhaps more strongly so than that exhibited by taxonomic clades (e.g. Gusmao et al. 2016, Norkko et al. 2019). Our findings on guilds as units of analysis are consistent with the general model of benthic trophic response to organic enrichment (Pearson & Rosenberg 1978, Gray et al. 2002), with richness, evenness, and composition all differing most strongly across space during Phase 1 (high local enrichment) and weakening significantly over Phases 2 and 3.

On the Palos Verdes shelf, the power of guild diversity was comparable to either family- or ordinal-level diversity, depending on the subset. The relatively high reliability of polychaete guild diversity as a proxy is possibly tied to how closely its functional diversity resembles that of the whole fauna (S/FG = 15.8 versus 14.9); the S/FG ratios of the less powerful bivalves and malacostracans were much lower (7.4

and 9, respectively). Because congeneric and confamilial species often encompass similar guilds (see Section 2), functional diversity might signal that a group will be a good proxy subset, even if genus- or family-level diversity is anomalously high.

Our observed increase in both guild richness and evenness with improved wastewater treatment (all 4 sets, although especially for bivalves; Figs. 4 & 5) is consistent with the decreases in functional diversity observed elsewhere under regimes of deteriorating conditions (e.g. Gusmao et al. 2016, Norkko et al. 2019). Future studies on taxonomic surrogacy should also evaluate guilds to consolidate the utility of this often-overlooked means of analysis, which, owing to conservatism within clades, can be achieved using relatively coarse taxonomic data.

We found that the ecological disparity of guilds, rather than their number alone, should be considered, that is, the range of life habit and trophic strategies represented. Bivalves, for example, span an exceptionally broad array of guilds, from deep-burrowing chemosymbiosis to free-swimming epifauna (both suspension feeders and predators), and thus can cover a broader environmental response along a pollution or other stress gradient than the larger number but more nuanced spectrum of guilds present in polychaetes and in malacostracans. Indeed, the functional response of bivalves to improved water treatment on the Palos Verdes shelf involves a precipitous decline in the abundance of chemosymbiont-bearing taxa in the late 1980s and the gradual increase of infaunal suspension and mixed feeders (Stull et al. 1996, LACSD 2012, Leonard-Pingel et al. 2019). A broad range of guilds may thus be a good signal that a taxonomic subset would be an effective indicator. A means of quantifying functional disparity among taxonomic subsets would help validate this potential method of subsetting taxonomic data.

5. CONCLUSIONS

This case study, using an exceptionally long-term and taxonomically consistent macrobenthic dataset from the urban Palos Verdes shelf of southern California, demonstrates that coarsening the taxonomic resolution of animal identification and focusing on taxonomic subsets of the whole fauna are effective techniques, individually and in combination, to reduce the taxonomic expertise and processing efforts of monitoring in regions with a strong suspected pollution gradient. Our findings that both polychaete and bivalve subsets detected weakening of the spa-

tial gradient over time, thus reflecting the whole fauna, and that patterns remained robust when coarsened to families or functional guilds, almost certainly reflect (1) the high strength of the original pollution gradient and (2) the preponderance of mono-specific genera, low-diversity families, and phylogenetic conservatism of traits that confer guild affinities and pollution tolerances of individual species in this fauna. These results thus strengthen the positive results for taxonomic sufficiency and caveats for taxonomic subsetting that have been emerging from tests elsewhere. Insights acquired from the highly resolved, methodologically standardized, and broad-scoped data produced by long-term monitoring programs in well-funded regions like southern California can be used to advise the design of efficient programs for areas where the fauna is less fully documented, local taxonomic expertise and/or funds are scarce, and yet the need for environmental evaluation is still or perhaps especially urgent.

LITERATURE CITED

Albano PG, Tomašových A, Stachowitsch M, Zuschin M (2016) Taxonomic sufficiency in a live–dead agreement study in a tropical setting. Palaeogeogr Palaeoclimatol Palaeoecol 449:341–348

Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. Austral Ecol 26:32–46

Bacci T, Trabucco B, Marzialetti S, Marusso V, Lomiri S, Vani D, Virno Lamberti C (2009) Taxonomic sufficiency in two case studies: Where does it work better? Mar Ecol 30(Suppl 1):13–19

Bevilacqua S, Fraschetti S, Musco L, Terlizzi A (2009) Taxonomic sufficiency in the detection of natural and human-induced changes in marine assemblages: a comparison of habitats and taxonomic groups. Mar Pollut Bull 58:1850–1859

Bevilacqua S, Terlizzi A, Claudet J, Fraschetti S, Boero F (2012) Taxonomic relatedness does not matter for species surrogacy in the assessment of community responses to

environmental drivers. J Appl Ecol 49:357–366

Bhusal DR, Kallimanis AS, Tsiafouli MA, Sgardelis SP (2014) Higher taxa vs. functional guilds vs. trophic groups as indicators of soil nematode diversity and community structure. Ecol Indic 41:25–29

Borja Á, Muxika I, Franco J (2006) Long-term recovery of soft-bottom benthos following urban and industrial sewage treatment in the Nervion estuary (southern Bay of Biscay). Mar Ecol Prog Ser 313:43–55

Borja Á, Chust G, Rodríguez J, Bald M and others (2016) 'The past is the future of the present': learning from long-time series of marine monitoring. Sci Total Environ 566-567:698–711

Bray JR, Curtis JT (1957) An ordination of upland forest communities of southern Wisconsin. Ecol Monogr 27:325–366

Brind'Amour A, Laffargue P, Morin J, Vaz S, Foveau A, Le Bris H (2014) Morphospecies and taxonomic sufficiency of benthic megafauna in scientific bottom trawl surveys. Cont Shelf Res 72:1–9

Buss DF, Vitorino AS (2010) Rapid bioassessment protocols using benthic macroinvertebrates in Brazil: evaluation of taxonomic sufficiency. J N Am Benthol Soc 29:562–571

Carneiro FM, Bini LM, Rodrigues LC (2010) Influence of taxonomic and numerical resolution on the analysis of temporal changes in phytoplankton communities. Ecol Indic 10:249–255

Caswell BA, Paine M, Frid CL (2018) Seafloor ecological functioning over two decades of organic enrichment. Mar Pollut Bull 136:212–229

Clare DS, Robinson LA, Frid CLJ (2015) Community variability and ecological functioning: 40 years of change in the North Sea benthos. Mar Environ Res 107:24–34

Clarke KR (1993) Non-parametric multivariate analyses of changes in community structure. Austral Ecol 18:117–143

Coan EV, Scott PV, Bernard FR (2000) Bivalve seashells of western North America. Santa Barbara Museum of Natural History Monographs 2. Santa Barbara Museum of Natural History, Santa Barbara, CA

Collins KS, Edie SM, Gao T, Bieler R, Jablonski D (2019) Spatial filters of function and phylogeny determine morphological disparity with latitude. PLOS ONE 14:e0221490

De Oliveira SS Jr, Ortega JCG, dos Santos Ribas LG, Lopes VG, Bini LM (2020) Higher taxa are sufficient to represent biodiversity patterns. Ecol Indic 111:105994

Di Minin E, Moilanen A (2014) Improving the surrogacy effectiveness of charismatic megafauna with well-surveyed taxonomic groups and habitat types. J Appl Ecol 51:281–288

Dimitriou PD, Apostolaki ET, Papageorgiou N, Reizopoulou S, Simboura N, Arvanitidis C, Karakassis I (2012) Meta-analysis of a large data set with Water Framework Directive indicators and calibration of a Benthic Quality Index at the family level. Ecol Indic 20:101–107

Eganhouse RP, Pontolillo J (2000) Depositional history of organic contaminants on the Palos Verdes Shelf, California. Mar Chem 70:317–338

Eleftheriou A, Moore DC (2013) Macrofauna techniques. In: Eleftheriou A (ed) Methods for the study of marine benthos. John Wiley & Sons, Chichester, p 175–251

Ellis D (1985) Taxonomic sufficiency in pollution assessment. Mar Pollut Bull 16:459

Ferraro SP, Cole FA (1992) Taxonomic level sufficient for assessing a moderate impact on macrobenthic communi-

ties in Puget Sound, Washington, USA. Can J Fish Aquat Sci 49:1184–1188

Fontaine A, Devillers R, Peres-Neto PR, Johnson LE (2015) Delineating marine ecological units: a novel approach for deciding which taxonomic group to use and which taxonomic resolution to choose. Divers Distrib 21: 1167–1180

Giangrande A, Licciano M, Musco L (2005) Polychaetes as environmental indicators revisited. Mar Pollut Bull 50: 1153–1162

Gogina M, Darr A, Zettler ML (2014) Approach to assess consequences of hypoxia disturbance events for benthic ecosystem functioning. J Mar Syst 129:203–213

Gray JS, Elliott M (2009) Ecology of marine sediments: from science to management, 2nd edn. Oxford University Press, New York, NY

Gray JS, Wu RS, Or YY (2002) Effects of hypoxia and organic enrichment on the coastal marine environment. Mar Ecol Prog Ser 238:249–279

Gusmao JB, Brauko KM, Eriksson BK, Lana PC (2016) Functional diversity of macrobenthic assemblages decreases in response to sewage discharges. Ecol Indic 66:65–75

Heino J (2010) Are indicator groups and cross-taxon congruence useful for predicting biodiversity in aquatic ecosystems? Ecol Indic 10:112–117

Hurlbert SH (1971) The nonconcept of species diversity: a critique and alternative parameters. Ecology 52:577–586

Islam MS, Tanaka M (2004) Impacts of pollution on coastal and marine ecosystems including coastal and marine fisheries and approach for management: a review and synthesis. Mar Pollut Bull 48:624–649

Jones FC (2008) Taxonomic sufficiency: the influence of taxonomic resolution on freshwater bioassessments using benthic macroinvertebrates. Environ Rev 16:45–69

Kotta J, Lauringson V, Kotta I (2007) Response of zoobenthic communities to changing eutrophication in the northern Baltic Sea. Hydrobiologia 580:97–108

LACSD (Los Angeles County Sanitation Districts) (2012) Joint Water Pollution Control Plant Biennial Receiving Water Monitoring Report 2010–2011. Los Angeles County Sanitation Districts, Whittier, CA

LACSD (2020) Joint Water Pollution Control Plant Biennial Receiving Water Monitoring Report 2018–2019. Los Angeles County Sanitation Districts, Whittier, CA

Landeiro VL, Bini LM, Costa FRC, Franklin E and others (2012) How far can we go in simplifying biomonitoring assessments? An integrated analysis of taxonomic surrogacy, taxonomic sufficiency and numerical resolution in a megadiverse region. Ecol Indic 23:366–373

Lee HJ, Sherwood CR, Drake DE, Edwards BD, Wong F, Hamer M (2002) Spatial and temporal distribution of contaminated, effluent-affected sediment on the Palos Verdes margin, southern California. Cont Shelf Res 22:859–880

Leonard-Pingel JS, Kidwell SM, Tomašových A, Alexander CR, Cadien DB (2019) Gauging benthic recovery from 20th century pollution on the southern California continental shelf using bivalves from sediment cores. Mar Ecol Prog Ser 615:101–119

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

Magierowski RH, Johnson CR (2006) Robustness of surrogates of biodiversity in marine benthic communities. Ecol Appl 16:2264–2275

Maurer D, Nguyen H, Robertson G, Gerlinger T (1999) The Infaunal Trophic Index (ITI): its suitability for marine environmental monitoring. Ecol Appl 9:699–713

Mellin C, Delean S, Caley J, Edgar G, Meekan M (2011) Effectiveness of biological surrogates for predicting patterns of marine biodiversity: a global meta-analysis. PLOS ONE 6:e20141

Moreno CE, Sánchez-Rojas G, Pineda E, Escobar F (2007) Shortcuts for biodiversity evaluation: a review of terminology and recommendations for the use of target groups, bioindicators and surrogates. Int J Environ Health 1: 71–86

Mueller M, Pander J, Geist J (2013) Taxonomic sufficiency in freshwater ecosystems: effects of taxonomic resolution, functional traits, and data transformation. Freshw Sci 32:762–778

Norkko J, Pilditch CA, Gammal J, Rosenberg R and others (2019) Ecosystem functioning along gradients of increasing hypoxia and changing soft-sediment community types. J Sea Res 153:101781

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

Olsgard F, Somerfield PJ (2000) Surrogates in marine benthic investigations — Which taxonomic unit to target? J Aquat Ecosyst Stress Recovery 7:25–42

Olsgard F, Brattegard T, Holthe T (2003) Polychaetes as surrogates for marine biodiversity: lower taxonomic resolution and indicator groups. Biodivers Conserv 12:1033–1049

Pearson TH, Rosenberg R (1978) Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. Oceanogr Mar Biol Annu Rev 16: 229–311

Pik AJ, Oliver I, Beattie AJ (1999) Taxonomic sufficiency in ecological studies of terrestrial invertebrates. Austral Ecol 24:555–562

Pitacco V, Mistri M, Aleffi IF, Lardicci C, Prato S, Tagliapietra D, Munari C (2019) The efficiency of taxonomic sufficiency for identification of spatial patterns at different scales in transitional waters. Mar Environ Res 144: 84–91

R Core Team (2020) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. www.r-project.org

Riedel B, Zuschin M, Stachowitsch M (2012) Tolerance of benthic macrofauna to hypoxia and anoxia in shallow coastal seas: a realistic scenario. Mar Ecol Prog Ser 458: 39–52

Ruppert KM, Kline RJ, Rahman MS (2019) Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: a systematic review in methods, monitoring, and applications of global eDNA. Glob Ecol Conserv 17:e00547

SCAMIT (Southern California Association of Marine Invertebrate Taxonomists) (2018) A taxonomic listing of benthic macro- and megainvertebrates from infaunal and epifaunal monitoring and research programs in the Southern California Bight, 12th edn. SCAMIT, Los Angeles, CA

Schiff K, Greenstein D, Dodder N, Gillett DJ (2016) Southern California Bight regional monitoring. Reg Stud Mar Sci 4:34–46

Schopf TJM (1978) Fossilization potential of an intertidal fauna: Friday Harbor, Washington. Paleobiology 4: 261–270

Schwing PT, O'Malley BJ, Romero IC, Martínez-Colón M and others (2017) Characterizing the variability of benthic foraminifera in the northeastern Gulf of Mexico following the Deepwater Horizon event (2010–2012). Environ Sci Pollut Res Int 24:2754–2769

Simonini R, Ansaloni I, Bonvicini Pagliai AM, Prevedelli D (2004) Organic enrichment and structure of the macrozoobenthic community in the northern Adriatic Sea in an area facing Adige and Po mouths. ICES J Mar Sci 61: 871–881

Smith RW, Bergen M, Weisberg SB, Cadien D and others (2001) Benthic response index for assessing infaunal communities on the southern California mainland shelf. Ecol Appl 11:1073–1087

Souza JLP, Baccaro FB, Landeiro VL, Franklin E, Magnusson WE, Pequeno PACL, Fernandes IO (2016) Taxonomic sufficiency and indicator taxa reduce sampling costs and increase monitoring effectiveness for ants. Divers Distrib 22:111–122

Stanley SM (1970) Relation of shell form to life habits of the Bivalvia (Mollusca). Geological Society of America Memoir 125. The Geological Society of America, Boulder, CO

Stark JS, Kim SL, Oliver JS (2014) Anthropogenic disturbance and biodiversity of marine benthic communities in Antarctica: a regional comparison. PLOS ONE 9: e98802

Stein ED, Cadien D (2009) Ecosystem response to regulatory and management actions: the southern California experience in long-term monitoring. Mar Pollut Bull 59:91–100

Stull JK, Swift DJ, Niedoroda AW (1996) Contaminant dispersal on the Palos Verdes continental margin. I. Sediments and biota near a major California wastewater discharge. Sci Total Environ 179:73–90

Swartz RC, Schults DW, Lamberson JO, Ozretich RJ, Stull JK (1991) Vertical profiles of toxicity, organic carbon, and chemical contaminants in sediment cores form the Palos Verdes Shelf and Santa Monica Bay, California. Mar Environ Res 31:215–225

Timms LL, Bowden JJ, Summersville KS, Buddle CM (2013) Does species-level resolution matter? Taxonomic sufficiency in terrestrial arthropod biodiversity studies. Insect Conserv Divers 6:453–462

Todd PA, Heery EC, Loke LHL, Thurstan RH, Kotze DJ, Swan C (2019) Towards an urban marine ecology: characterizing the drivers, patterns and processes of marine ecosystems in coastal cities. Oikos 128:1215–1242

Tyler CL, Kowalewski M (2017) Surrogate taxa and fossils as reliable proxies of spatial biodiversity patterns in marine benthic communities. Proc R Soc B 284:20162839

Underwood AJ (1991) Beyond BACI: experimental designs for detecting human environmental impacts on temporal variations in natural populations. Mar Freshw Res 42:569–587

Underwood AJ (1994) On Beyond BACI: sampling designs that might reliably detect environmental disturbances. Ecol Appl 4:3–15

Villnäs A, Norkko J, Lukkari K, Hewitt J, Norkko A (2012) Consequences of increasing hypoxic disturbance on benthic communities and ecosystem functioning. PLOS ONE 7:e44920

Warwick RM (1988) The level of taxonomic discrimination required to detect pollution effects on marine benthic communities. Mar Pollut Bull 19:259–268

Warwick RM (1993) Environmental impact studies on marine communities: pragmatic considerations. Austral Ecol 18:63–80

Westgate MJ, Barton PS, Lane PW, Lindenmayer DB (2014) Global meta-analysis reveals low consistency of biodiversity congruence relationships. Nat Commun 5:3899

Włodarska-Kowalczuk M, Kędra M (2007) Surrogacy in natural patterns of benthic distribution and diversity: selected taxa versus lower taxonomic resolution. Mar Ecol Prog Ser 351:53–63