



NOTE

Taxonomic surrogates for long-term macrobenthic community monitoring: an application with community trajectory analysis

Chirine Toumi^{1,*}, Olivier Gauthier^{1,2}, Éric Thiébaud^{3,4}, Clément Guedes¹, Jacques Grall^{1,2}

¹LEMAR, Univ Brest, CNRS, IRD, Ifremer, 29280 Plouzané, France

²OSU IUEM, Univ Brest, CNRS, IRD, 29280 Plouzané, France

³Sorbonne Université, CNRS, Station Biologique de Roscoff, UMR7144, Adaptation et Diversité en Milieu Marin, Place Georges Teissier, CS90074, 29688 Roscoff Cedex, France

⁴Sorbonne Université, CNRS, OSU STAMAR, UAR2017, 4 Place Jussieu, 75252 Paris Cedex 05, France

ABSTRACT: Biodiversity monitoring, essential to detect impacts of natural and anthropogenic changes on marine ecosystems, is costly, time-consuming and requires high taxonomic expertise. Taxonomic surrogacy might be a solution to overcome these problems and accurately reflect species-level community patterns, but its efficiency has mainly been assessed as taxonomic sufficiency and rarely from long-term monitoring data. Here, the efficiency of subset taxa (i.e. Polychaeta, Crustacea and Mollusca) for summarizing long-term community dynamics was tested in different coastal habitats. The data set came from a yearly long-term macrobenthic monitoring programme (2007–2019) in western France, in 2 biogenic and 2 bare habitats. Community trajectory analysis (CTA), a statistical approach allowing for quantitative measures and comparisons of temporal trajectories, was used to test for similarities between overall community, subset-taxa and non-subset-taxa dynamics. Polychaeta best reflected the spatial diversity of the different sites as well as the temporal dynamics of the non-Polychaeta species, with more efficiency in biogenic compared to bare habitats. Our study confirmed that the subset-taxon method may reflect long-term benthic habitat dynamics and that CTA is an effective tool to test their efficiency.

KEY WORDS: Taxonomic surrogacy · Community trajectory analysis · Benthic macrofauna · Monitoring

1. INTRODUCTION

Long-term biodiversity monitoring is essential to detect changes in ecosystem structure and functioning (Dornelas et al. 2013). Biodiversity is mostly monitored through species-level morphological identification of organisms, a costly and time-consuming

task requiring high taxonomic expertise (Olsgard & Somerfield 2000, Wodarska-Kowalczyk & Kedra 2007). Biodiversity surrogates can help overcome these difficulties; more readily estimated, they might strongly correlate with species richness and reflect species-level community patterns consistently (Olsgard & Somerfield 2000). Different types of surro-

*Corresponding author: chirine.toumi@univ-brest.fr

gates exist: a higher taxon as a surrogate for a lower taxonomic level (higher-taxon method), a taxon for another of the same taxonomic level (cross-taxa method) and a taxon for the entire target community (subset-taxon method) (Mellin et al. 2011).

For soft-bottom macrofauna, commonly used to assess anthropogenic impacts or environmental changes in coastal marine ecosystems, the relevance of taxonomic surrogacy to address changes in community composition and structure has already been evaluated, frequently over gradients of pollution or disturbances (Olsgard & Somerfield 2000, Kokesh et al. 2022). However, most studies used the higher-taxon method rather than other surrogates (Bevilacqua et al. 2012), and scarcely in the context of long-term monitoring (but see Pitacco et al. 2019). The few studies focusing on long-term series relied on methods not explicitly designed for temporal dynamics (e.g. Kokesh et al. 2022). The most investigated subset taxa were Polychaeta, Crustacea, Mollusca and Echinodermata, which are generally the most diverse and abundant taxonomic groups in soft-bottom communities. Polychaeta were frequently found to be the most reliable subset taxon (Olsgard et al. 2003, Kokesh et al. 2022). However, for long-term monitoring, little attention has been given to the assumption that surrogates of taxonomic diversity are constant over time (Magierowski & Johnson 2006), a statement that needs to be quantitatively assessed. Finally, the efficiency of the subset-taxon method may be habitat-dependent and may vary with type and magnitude of disturbances (Wodarska-Kowalczyk & Kedra 2007, Mellin et al. 2011).

While many methods have been developed to describe long-term evolution of marine communities, community trajectory analysis (CTA) is a recently developed multivariate method specifically tailored to study temporal community dynamics (De Cáceres et al. 2019). From a classical ordination of a species abundance matrix, it performs a geometric analysis of temporal trajectories in all dimensions to characterize and compare temporal patterns in community dynamics.

Using CTA and data from 13 yr of monitoring benthic macrofauna at 32 sites along the coast of Brittany (France) (<https://rebenit.ifremer.fr/>), we investigated whether long-term dynamics of soft-bottom communities could be summarized by the subset-taxon method and which subset taxon among Polychaeta, Crustacea and Mollusca was the most efficient. Subset-taxa temporal trajectories were compared with those of overall and non-subset-taxa communities (i.e. whole community minus subset taxon). We ad-

ressed these questions using 4 soft-bottom habitats exposed to different environmental constraints: 2 associated with foundation species, i.e. eelgrass and maerl beds, and 2 bare sedimentary habitats. Focusing on spatial and temporal changes in α and β taxonomic diversities, we hypothesized that: (1) surrogate performance for community dynamics is taxon-dependent and (2) the efficiency of the subset-taxon method is habitat-dependent.

2. MATERIALS AND METHODS

2.1. Sampling

Macrobenthic species abundances from 32 sites were recorded yearly from 2007 to 2019 in 4 habitats along 500 km of the coast of Brittany (France): intertidal eelgrass meadows (8 sites), intertidal sandy beaches (9 sites), subtidal maerl beds (7 sites) and subtidal soft sediments (8 sites), respectively referred to as intertidal biogenic (IBIO), intertidal bare (IBAR), subtidal biogenic (SBIO) and subtidal bare (SBAR) habitats. We collected 9–10 replicates at each site (intertidal: 0.03 m² core, subtidal: 0.1 m² Smith-McIntyre grab; see Fig. S1 in the Supplement at www.int-res.com/articles/suppl/m714p105_supp.pdf). Specimens were identified to the lowest possible taxonomic level (mostly species), and taxonomic homogenization was performed to ensure consistent taxonomic resolution over time and space (Text S1).

2.2. Data analysis

The relationships between the number of subset-only (i.e. species in the subset) and non-subset (i.e. the remaining species) species was assessed with ordinary least squares regressions and tested using 999 permutations (Legendre & Legendre 2012).

We used Hellinger distances to analyse β -diversity between sites and years for: (1) overall, (2) non-subset and (3) subset-only communities in the 4 habitats. Observations were represented with principal coordinates analysis (PCoA), and consecutive observations of a given site were linked by a segment; all segments of a site constitute its temporal trajectory. CTA (De Cáceres et al. 2019) was then performed on all dimensions. Resemblance between trajectory pairs was assessed with symmetrized directed segment path dissimilarity (D_{SDSP}) (De Cáceres et al. 2019) using their geometry (shape, size, direction, position). Because trajectory position might be influ-

enced by site-specific factors, we centred trajectories prior to D_{SDSP} calculation to focus on compositional dynamics rather than spatial variation.

Co-inertia analysis seeks common structures between data sets, and the RV coefficient, a multivariate generalization of the squared Pearson correlation, measures the closeness of 2 separate ordinations (Legendre & Legendre 2012). It ranges from 0 to 1 and was tested with 999 permutations. We tested the co-structure between non-subset and subset-only data sets as the whole community, and subset-only data sets are not independent. RVs were computed for 3 different ordinations: PCoA representing Hellinger distances between observations (raw trajectories), PCoA representing centred trajectories and PCoA representing dissimilarity of site dynamics (D_{SDSP}). The first configuration takes into account spatial differences in community composition, the second smooths these differences to focus on temporal dynamics, and the last explicitly compares similarities between community dynamics using a CTA metric. All dimensions of PCoAs were considered to compute RV coefficients.

All analyses were conducted with the R programming language version 4.2.2 (R Core Team 2022) with the packages 'ecotraj' (De Cáceres et al. 2019, Sturbois et al. 2021) and 'ade4' (Thioulouse et al. 2018).

3. RESULTS

Results revealed significant positive linear relationships between the number of non-subset and subset-only species in all habitats except for Crustacea in IBIO (Fig. 1), with the highest R^2 between non-Polychaeta and Polychaeta-only in IBAR and the lowest for non-Polychaeta and Polychaeta-only in IBIO. Crustacea were absent from several IBAR samples, however their total richness and abundance was sometimes higher than those of Polychaeta (Table 1).

Fig. 2 represents 2-dimensional community trajectories of IBIO as an example. Total variance of the first 2 PCoA axes ranged from 41% (non-Polychaeta) to 52% (Polychaeta-only). Visually considering overall community trajectories, sites clearly occupied different positions in the biplot, reflecting site-specific composition and structure. Site trajectory positions were more similar between overall community and Polychaeta-only compared to Mollusca-only or Crustacea-only. Moreover, non-Polychaeta trajectory positions diverged more from those of the overall

community than either non-Mollusca or non-Crustacea. Polychaeta might perform better when representing spatial β -diversity patterns than Crustacea or Mollusca and might drive the overall community dynamics. Such considerations were mainly true for every studied habitat (Figs. S2–S4).

RVs of raw trajectories were all significant (Table 2). The highest values were found between non-Polychaeta and Polychaeta-only in all habitats, supporting the idea that Polychaeta were better surrogates for studying β -diversity between sites and years. After trajectory centering, RV coefficients all strongly decreased but remained significant. Interestingly, ordinations of centred trajectories were still the closest for Polychaeta. Most RV coefficients on PCoAs based on D_{SDSP} were not significant. In this configuration, significant RVs mean that 2 sites with similar dynamics in the subset-only data set have similar dynamics in the non-subset data set. This was stronger for Polychaeta in biogenic habitats. Indeed, comparing non-Polychaeta and Polychaeta-only data sets, RVs were significant in IBIO and SBIO. In contrast, for non-Mollusca and Mollusca-only data sets, RVs were never significant, while non-Crustacea and Crustacea-only data sets showed significant RV only in SBIO.

4. DISCUSSION

Our study covers a diversity of habitats and years to generalize results, alleviating difficulties or biases potentially associated with meta-analyses conducted from data sets with highly contrasting environments and sampling strategies. Similar to previous local studies, our results showed that Polychaeta better translated both community spatial differences and temporal dynamics compared to Crustacea or Mollusca (e.g. Olsgard & Somerfield 2000, Wodarska-Kowalczyk & Kedra 2007, Kokesh et al. 2022). However, in contrast with most studies on temporal dynamics which used non-metric multidimensional scaling (nMDS), sometimes combined with Mantel tests (Gladstone et al. 2020), a method that has been subject to criticisms (Legendre et al. 2015), we assessed the effectiveness of taxonomic surrogacy with a method explicitly designed to quantitatively evaluate temporal dynamics.

Polychaeta efficiency could scarcely be linked to their numerical dominance or richness, as they were sometimes surpassed by Crustacea in both (Table 1). However, Polychaeta do harbour a high diversity of biological traits (Olsgard et al. 2003), with a very wide range of feeding, motility and reproduction

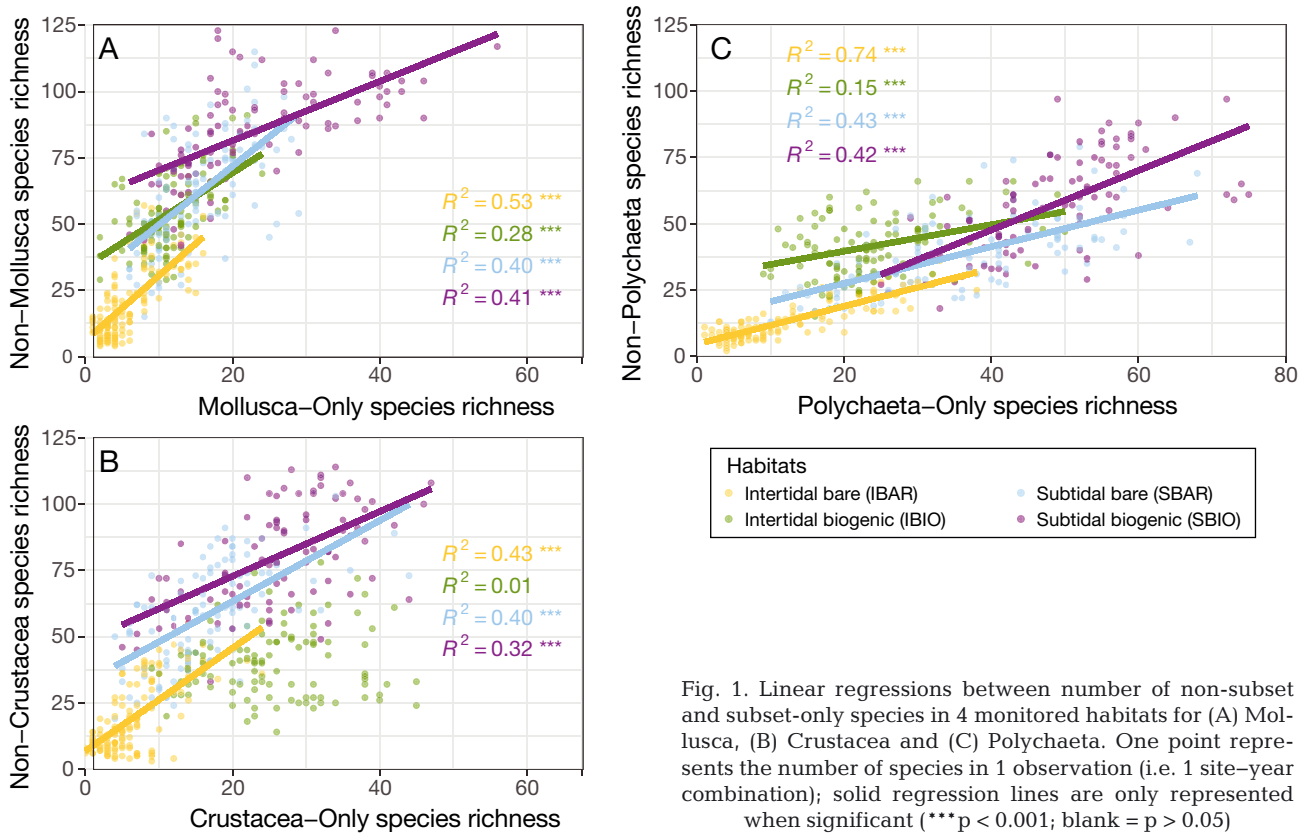


Fig. 1. Linear regressions between number of non-subset and subset-only species in 4 monitored habitats for (A) Mollusca, (B) Crustacea and (C) Polychaeta. One point represents the number of species in 1 observation (i.e. 1 site-year combination); solid regression lines are only represented when significant (*** $p < 0.001$; blank = $p > 0.05$)

Table 1. Total and relative species richness and abundance of each subset in each habitat: intertidal bare (IBAR), intertidal biogenic (IBIO), subtidal bare (SBAR) and subtidal biogenic (SBIO)

Subsets		IBAR	IBIO	SBAR	SBIO
Species richness	Overall community	332	493	514	648
	Mollusca	71	93	119	151
	Crustacea	109	184	160	178
	Polychaeta	124	156	185	250
Species richness (%)	Mollusca	21	19	23	23
	Crustacea	32	37	31	27
	Polychaeta	37	32	36	39
Total abundance	Overall community	47 365	128 424	129 007	237 463
	Mollusca	6812	10 557	27 841	25 057
	Crustacea	23 510	45 482	19 994	147 663
	Polychaeta	16 175	67 795	70 880	56 939
Total abundance (%)	Mollusca	14	8	22	11
	Crustacea	49	35	15	62
	Polychaeta	34	52	55	24

modes and life spans (Giangrande 1997, Jumars et al. 2015). They also include species sensitive and tolerant to disturbances (Olsgard et al. 2003, del Pilar Ruso et al. 2009). Thus, we hypothesize that Polychaeta mimicked the range of functional guilds encountered in the whole community, hence the dynamics of a majority of non-Polychaeta taxa. Their

high functional diversity would also allow dealing with various environmental conditions and/or habitats (Jumars et al. 2015), which is convenient for studies involving environmental gradients or different habitats.

Although Polychaeta performed better, their efficiency as a subset-taxon was habitat-dependent.

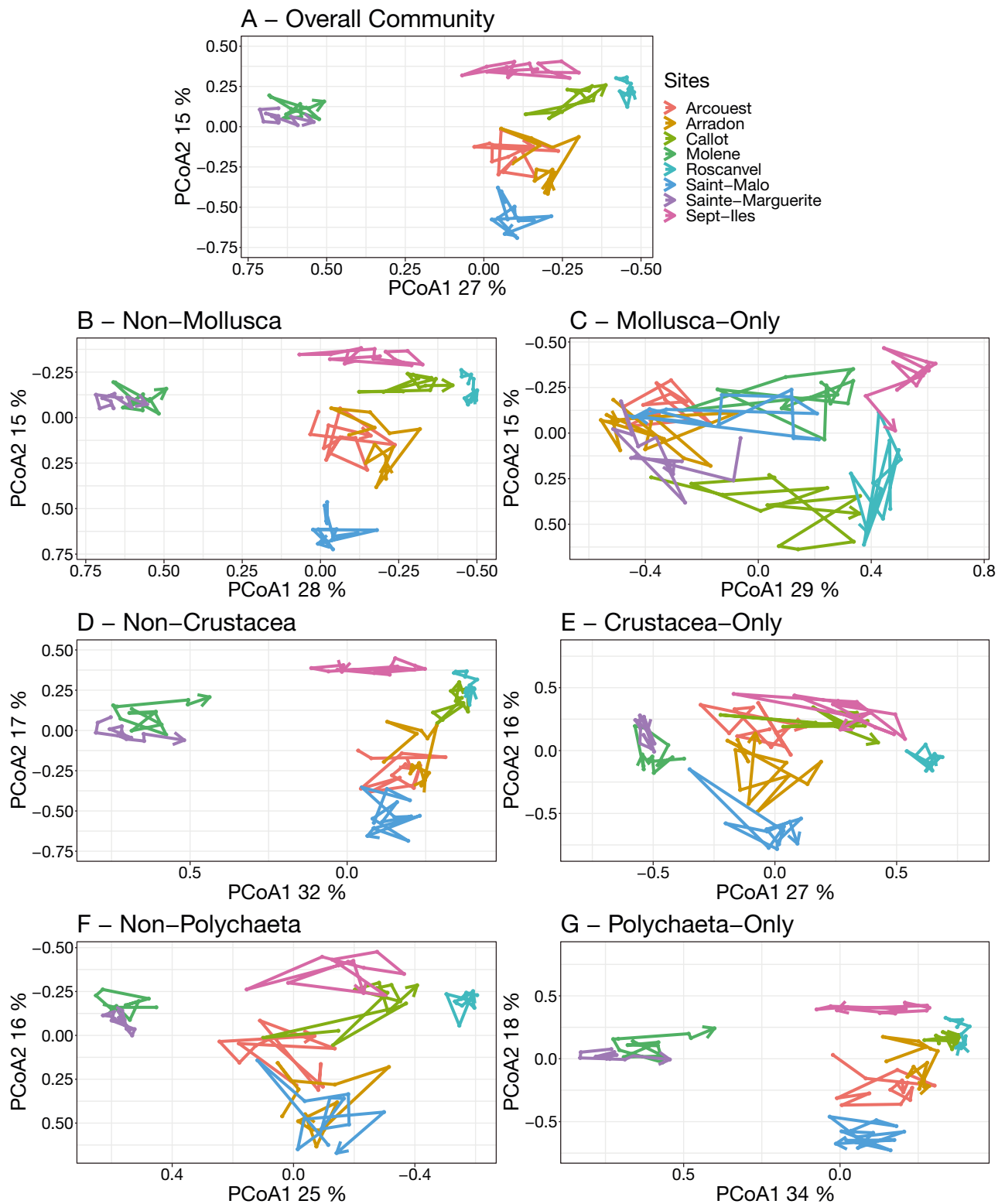


Fig. 2. Two-dimensional (principal coordinates analysis, PCoA) representation of community trajectories (2007–2019) of the (A) overall community, (B,D,F) non-subset and (C,E,G) subset-only in the intertidal biogenic habitat. One point represents the community state of a site in a given year (1 observation). Site-specific consecutive community states are linked by segments, which taken together depict the site trajectory. Arrows represent the final community state of a trajectory. Site locations are shown in Fig. S1 in the Supplement

Table 2. RV coefficients between configurations of points in 2 ordinations (always between the ordination of the non-subset data set under consideration, i.e. Mollusca, Crustacea and Polychaeta, and the subset-only ordination) in intertidal bare (IBAR), intertidal biogenic (IBIO), subtidal bare (SBAR) and subtidal biogenic (SBIO) habitats. Principal coordinate analysis (PCoA) ordinations represent different point configurations; raw trajectories: consecutive observations in a PCoA based on their Hellinger distances and linked by segments; centered trajectories: previous ordination after centring trajectories; and D_{SDSP} : PCoA of centred trajectories using symmetrized directed segment path dissimilarity (here a point represents 1 site trajectory). Significance is indicated by asterisks: * $p < 0.05$, *** $p < 0.001$ (blank = $p > 0.05$). Schemes represent ordinations of 2 axes as in Fig. 2, but RV coefficients were computed on all dimensions

Configuration	Subset	IBAR	IBIO	SBAR	SBIO
Raw trajectories					
	Mollusca	0.67***	0.51***	0.77***	0.68***
	Crustacea	0.67***	0.74***	0.44***	0.71***
	Polychaeta	0.77***	0.81***	0.80***	0.75***
Centered trajectories					
	Mollusca	0.22*	0.29***	0.43***	0.42***
	Crustacea	0.23***	0.40***	0.44***	0.42***
	Polychaeta	0.26***	0.42***	0.47***	0.48***
D_{SDSP}					
	Mollusca	0.73	0.94	0.95	0.93
	Crustacea	0.90	0.96	0.94	0.93*
	Polychaeta	0.89	0.97*	0.96	0.96*

Such discrepancies have already been shown (Maggierowski & Johnson 2006, Wodarska-Kowalczyk & Kedra 2007, Mellin et al. 2011), emphasizing the need to choose surrogates according to target habitat or community. As such, Polychaeta were the best predictors of α -diversity in all habitats except IBIO, where Mollusca performed better. Nevertheless, Polychaeta were still the best predictor of spatial β -diversity and community dynamics in IBIO. Indeed, positive correlations between site dynamic dissimilarities of Polychaeta-only and non-Polychaeta were found in both biogenic habitats (IBIO and SBIO). The high complexity of maerl beds (SBIO) may enhance niche partitioning, leading to high species diversity and functional redundancy in this habitat, that evenly harbours all modalities of Polychaeta traits at each site (Boyé et al. 2019). In this habitat, the high functional diversity of Polychaeta may allow for a good representation of the dynamics of other taxa with similar ecological functions. However, this

might not hold for IBIO, where Polychaeta were less numerous and displayed specialized traits (Boyé et al. 2019) which are unlikely to mimic the overall community. An alternative hypothesis could be that biogenic habitats enhance stability (Toumi et al. 2023), allowing for the maintenance of the same dynamic for Polychaeta-only and non-Polychaeta.

We expect more fluctuant dynamics in bare habitats, especially the more exposed IBAR (Toumi et al. 2023). Bare habitats also harbour less species diversity for a given tidal zone, and Polychaeta in IBAR have site-specific biological traits (Boyé et al. 2019). Less numerous taxa showing various responses to frequent stress may explain the reduced efficiency in bare habitats. Higher variability of bare environments could lead to greater stochasticity in community dynamics, hindering establishing a link between surrogates and community dynamics.

Although correlations between site dynamic dissimilarities were only significant in biogenic habitats

for Polychaeta, we cannot exclude their efficiency in bare habitats: considering raw or centred configurations, RVs were always significant and maximised for Polychaeta. These configurations reflect taxonomic temporal dynamics; thus we posit that they were translated in bare habitats, and Polychaeta were the best subset to do so.

Significant RV values were highest for raw configurations, reinforcing the suitability of Polychaeta as spatial β -diversity surrogates. Centred configurations revealed that β -diversity was more spatial than temporal. RV values were highest but non-significant with D_{SDSP} , but in this instance, power is reduced as sample size is the number of sites (maximum 9 in this study).

CTA appears as a suitable tool to quantitatively compare dynamics between community subsets. Polychaeta were the best subset-taxon for long-term dynamics of soft-bottom macrobenthic diversity. Their efficiency was demonstrated in all habitats, especially biogenic ones. Prior to initiating a monitoring programme based on a community subset, we recommend testing the efficiency of taxonomic surrogacy on a time series of the entire community if available, or to refer to surrogacy studies on the same kind of environment beforehand.

Acknowledgements. We thank the crews of the RVs 'Thalia', 'Côtes de la Manche', 'Neomysis' and 'Albert Lucas' for their assistance during field surveys. We gratefully acknowledge Caroline Broudin, Céline Houbin, Marion Maguer, Vincent Le Garrec, and all students, technicians and colleagues who contributed to sampling and laboratory work since 2007. This study was supported by the REBENT programme coordinated by Sandrine Derrien (MNHN) and its funding partners (Agence de l'Eau Loire Bretagne, Région Bretagne, DREAL Bretagne) and by IUEM. We thank Thomas Burel for his help with figures.

LITERATURE CITED

- 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
- Boyé A, Thiébaud É, Grall J, Legendre P and others (2019) Trait-based approach to monitoring marine benthic data along 500 km of coastline. *Divers Distrib* 25: 1879–1896
- De Cáceres M, Coll L, Legendre P, Allen RB and others (2019) Trajectory analysis in community ecology. *Ecol Monogr* 89:e01350
- Del Pilar Ruso Y, De la Ossa Carretero JA, Loya Fernández A, Ferrero Vicente LM, Giménez Casalduero F, Sánchez Lizaso JL (2009) Assessment of soft-bottom Polychaeta assemblage affected by a spatial confluence of impacts: sewage and brine discharges. *Mar Pollut Bull* 58: 776–782
- Dornelas M, Magurran AE, Buckland ST, Chao A and others (2013) Quantifying temporal change in biodiversity: challenges and opportunities. *Proc R Soc B* 280: 20121931
- Giangrande A (1997) Polychaete reproductive patterns, life cycles and life histories: an overview. *Oceanogr Mar Biol Annu Rev* 35:323–386
- Gladstone W, Murray BR, Hutchings P (2020) Promising yet variable performance of cross-taxon biodiversity surrogates: a test in two marine habitats at multiple times. *Biodivers Conserv* 29:3067–3089
- Jumars PA, Dorgan KM, Lindsay SM (2015) Diet of worms emended: an update of polychaete feeding guilds. *Annu Rev Mar Sci* 7:497–520
- Kokesh BS, Kidwell SM, Tomašových A, Walther SM (2022) Detecting strong spatial and temporal variation in macrobenthic composition on an urban shelf using taxonomic surrogates. *Mar Ecol Prog Ser* 682:13–30
- Legendre P, Legendre LFJ (2012) Numerical ecology, 3rd edn. *Developments in environmental modelling*, Vol 24. Elsevier, Oxford
- Legendre P, Fortin MJ, Borcard D (2015) Should the Mantel test be used in spatial analysis? *Methods Ecol Evol* 6: 1239–1247
- Magierowski RH, Johnson CR (2006) Robustness of surrogates of biodiversity in marine benthic communities. *Ecol Appl* 16:2264–2275
- Mellin C, Delean S, Caley J, Edgar G and others (2011) Effectiveness of biological surrogates for predicting patterns of marine biodiversity: a global meta-analysis. *PLOS ONE* 6:e20141
- 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
- Pitacco V, Mistri M, Infantini V, Sfriso A, Sfriso AA, Munari C (2019) Benthic studies in LTER sites: the use of taxonomy surrogates in the detection of long-term changes in lagoonal benthic assemblages. *Nat Conserv* 34:247–272
- R Core Team (2022) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Sturbois A, De Cáceres M, Sánchez-Pinillos M, Schaal G and others (2021) Extending community trajectory analysis: new metrics and representation. *Ecol Model* 440: 109400
- Thioulouse J, Dray S, Dufour AB, Siberchicot A, Jombart T, Pavoine S (2018) Multivariate analysis of ecological data with ade4. Springer, New York, NY
- Toumi C, De Cáceres M, Grall J, Boyé A and others (2023) Long-term coastal macrobenthic community trajectory analysis reveals habitat-dependent stability patterns. *Ecography* 2023:e06489
- Wodarska-Kowalczyk M, Kedra M (2007) Surrogacy in natural patterns of benthic distribution and diversity: selected taxa versus lower taxonomic resolution. *Mar Ecol Prog Ser* 351:53–63