ORIGINAL ARTICLE

Taxonomic sufficiency in two case studies: where does it work better?

Tiziano Bacci, Benedetta Trabucco, Sara Marzialetti, Veronica Marusso, Serena Lomiri, Danilo Vani & Claudia Virno Lamberti

ISPRA, Formerly ICRAM, Central Institute for Marine Research, Rome, Italy

Abstract

In marine macrobenthos studies the identification of organisms at species level is the best entry to ecological and biological information about the animals. An accurate identification requires excellent conservation of the organisms, reliable fauna description, experts and lengthy work in the laboratory. The aim of this work is to test taxonomic sufficiency (TS) in two deliberately selected different case studies to understand whether and how the taxonomic complexity of a benthic assemblage influences the results of TS and where it works better. The first benthic settlement was collected in an area characterized by homogeneous depth and grain size composition (case study A) around an off-shore gas platform, while the second one was collected along a coast-wide transect in an area with human pressure limited to fishing activities (case study B). Univariate and multivariate statistical analysis was used to assess differences in the taxonomic structure of benthic assemblages and to test TS on the two different datasets. TS seems to work in both sites, from species to higher taxonomic levels, and the family taxonomic level appears the best compromise for taxonomic resolution when an accurate identification is not achievable. The application of TS does not indicate a significant difference between the two datasets and appears therefore to be a valid instrument to analyse and describe the structure of benthic settlements in the case of taxonomically complex communities.

Problem

The importance of soft-bottom macrofauna in environmental and ecological studies arises from several factors: sedimentary habitats cover most of the ocean bottom and constitute the largest single ecosystem on earth in terms of spatial coverage (Snelgrove 1997); most marine species are benthic (Gray 1997) and have the ability to respond significantly to natural and anthropogenic environmental variations (Pearson & Rosenberg 1978; Gray 1981; Gray et al. 1990; Warwick & Clarke 1991; Castelli et al. 2003). Anthropogenic impact studies on the marine environment often use benthic macrofaunal assemblages as indicators: these organisms respond promptly to environmental disturbances owing to their relatively short lifespans, and

they are small and can be sampled quantitatively (Warwick 1993; Vanderklift et al. 1996).

With reference to the role of systematics in assessing pollution effects, Carriker (1976) stressed the importance of accurate identification to unlock the storage and retrieval system of scientific information, given that a species name is the doorway to its literary pedigree and makes information accessible and useful via published aspects of organism biology.

Unfortunately, an accurate identification often requires an excellent conservation of the organisms, the presence of a reliable fauna description, more experts and lengthy work in the laboratory. The easiest way to simplify this process would be to use coarser taxonomic resolutions (e.g. genus instead of species or family instead of genus).

Taxonomic sufficiency (TS) involves identifying taxa to the highest category possible (genus, family, order, *etc.*) without losing statistically significant vigour in assessing impacts (Ellis 1985).

Literature on the application of TS is abundant and is spread over different geographical areas and different kinds of impact studies, such as oil extraction fields (Olsgard *et al.* 1997), oil spill (Gomez Gesteira *et al.* 2003), heavy metal pollution (Vanderklift *et al.* 1996), and organic enrichment (Castanedo *et al.* 2007), using diverse sampling procedures (Ferraro & Cole 1992), the relationship between taxonomic resolution and spatial scales (Anderson *et al.* 2005) and uses of datasets (Warwick 1988).

Most of these studies (Dauvin *et al.* 2003; Raymond *et al.* 2005) suggest that a lower taxonomic resolution can be sufficient when studying benthic assemblages composition and that determining the family may be satisfactory in many routine monitoring surveys. In highly disturbed areas, it is suggested that the results of multivariate analyses based on higher taxa might more closely reflect gradients of contamination or stress than those based on species data, the latter being more affected by natural environmental variables.

However, not all researchers agree with the use of TS-based approaches and several authors (Dauvin *et al.* 2003; Raymond *et al.* 2005) stress that when studying poorly known environments, it is essential to acquire detailed information before applying a rougher taxonomic resolution; therefore, the use of TS is likely to be the most advantageous if species level baseline studies have already been completed. Terlizzi *et al.* (2003) emphasizes that evidence suggests that family is a sufficient taxonomic level, but underlines that this statement stems from too limited a number of case studies. Relations among TS and sampling procedures, data analyses, spatial scale, habitat features and assemblages structure remain to be assessed and are far from being generalized.

Objectives Statement

This work intervenes in the debate on TS application as a contribution towards a further understanding of the diverse possibilities for using this instrument in diverse datasets. Our aim is to test TS in two deliberately selected different case studies to understand whether and how the structure and the taxonomic complexity of a benthic assemblage influences the results of taxonomic sufficiency, and where it works better.

The first study concerned animals collected around an off-shore gas platform at a homogeneous depth, where granulometric gradient is absent. This benthic assemblage was characterized by a small number of species and a simple taxonomic structure. A second dataset referred to a settlement collected along a coast-wide transect with anthropogenic pressure limited to fishing activities (G. Franceschini, personal communication). In this case the benthic assemblage was characterized by a high number of species and a complex taxonomic structure.

Methods

Case study A was related to the monitoring of an off-shore gas platform located in the Central Adriatic Sea sited at a depth of about 120 m. A radial sampling design of 12 stations positioned at increasing distances from the installation (0, 50, 500, 1000, 2000 m) was developed to evaluate the spatial distribution of benthic assemblages and potential effects induced by the presence of the platform and its influence on them (Trabucco *et al.* 2006). Station A0 was located at the installation, stations A1, A2, A3, A4 at 50 m from it, stations A5, A6, A7, A8 at 500 m from it, station A9 at 1000 m from it, and control stations K2 and K3 were positioned at the upper-current about 2000 m NW of the platforms, on a bottom which presented geo-morphological characteristics similar to the study area.

Case study B was developed to monitor an area along a coast-wide transect in the North Adriatic Sea. The sampling was carried out at 33 stations on a bathymetric gradient from 1 to 29 m depth.

Data analyses were done on data pooled from two samples collected at each station using a Van Veen grab (0.1 m²; 25 l) and processed through a sieve (1-mm mesh size). The macrozoobenthic component was sorted and identified to the lowest possible taxonomic level, *i.e.* species level.

Species abundances of Polychaeta, Mollusca, Crustacea and Echinodermata from each dataset were aggregated at genus, family, order, class and phylum levels, following the MARBEF database taxonomic classification (http://www.marbef.org/data/erms.php).

At each station, variations in the taxonomic structure of benthic assemblages were analyzed through the difference $NT_x - NT_{x+1}$ defined as 'loss of information α' (α). NT expresses the number of taxa identified at the station at a certain taxonomic level X. X can range from species to phylum level. X + 1 expresses instead the next higher taxonomic level compared to the one which is taken into consideration as X (*i.e.* if the X level is species, X + 1 will be genus). The parameter 'α' describes the loss of information on the taxonomic structure of a benthic settlement which can be registered at the passage from a lower (X) to a higher (X + 1) taxonomic level. 'α' values therefore give general information about the taxonomic heterogeneity-complexity inside each level considered. Higher

values correspond to a major 'loss information α' and *vice versa*, while a value equal to zero explains a null 'loss information α'. In this study the 'α' was calculated step by step from the lowest to the highest taxonomic level considered (species-genus, genus-family, family-order, order-class, class-phylum).

Statistical analysis was performed using the PRIMER package, version 6 (Clarke & Gorley 2006). Assemblage abundance matrices were produced for each of the six taxonomic levels (species, genus, family, order, class, phylum) and were square-root transformed to reduce the weighting of the most abundant taxa.

The affinities among the stations were established using non-metric Multi-Dimensional Scaling (MDS) calculated with 'zero-adjusted' Bray–Curtis similarity (Bray & Curtis 1957; Clarke *et al.* 2006) using Kruskal's stress Formula 1 (Kruskal & Wish 1978; Clarke & Green 1988). The stress value (S) measures the degree of coupling of real between-sample distances and between-sample distances in the ordination: high stress values thus indicate poor reliability, as apparent differences between samples may simply be numerical artefacts (Gomez Gesteira *et al.* 2003). As a rough guide to efficiency, $S > 0.5$ is probably random, $0.5 > S > 0.25$ is a poor result, $0.25 > S > 0.1$ is a satisfactory result and $S < 0.1$ is a good result (Kruskal 1964). Moreover, hierarchical clustering (group average) on species, genus, family, order, class and phylum taxonomic levels was calculated and visualized in the MDS ordination plots. A one-way ANOSIM test (Clarke & Green 1988) was performed among the groups that stood out at species level analysis at each analyzed taxonomic level. Finally, Spearman's rank correlation coefficient (ρ) was calculated between matrices at species level and at higher taxonomic levels.

Results

Case study A

Dataset A comprises a total of 287 individuals from the phyla Polychaeta, Mollusca, Crustacea and Echinodermata. Forty-nine taxa were identified at species level.

The benthic assemblage shows a different species composition in the stations near the platform (located 0–50 m from it, 'inner group'), in comparison with the stations located at greater distances (500, 1000, 2000 m, 'outer group'). The presence of such differences between the 'inner' and 'outer' group is related to the platform structure, which promotes the presence of a coarser secondary substratum and therefore the occurrence of different species (Trabucco *et al.* 2006).

A low number of species and individuals and a simple taxonomic structure were recorded. Only two genera

contained two species, the rest contained only one. Similarly, most families contained only one species: families with two species were uncommon and only one family contained four species. 'α' values confirm that the taxonomic structure of the benthic assemblage is generally very simple, especially in the 'outer group' stations (Fig. 1). In the 'inner group', two stations (A1, A3) show a 'loss of information α' descending from species to genus level because there are one or more genera which comprise more than one species. Four stations (A0, A1, A3, A4) show loss of information between genus and family levels, indicating the presence of one or more families with more than one genus. Five stations (A0, A1, A2, A3, A4) show loss at family-order level because of the presence of one or more orders with more than one family. Five stations (A0, A1, A2, A3, A4) show loss at order-class level due to the presence of one or more classes with more than one order. Finally, two stations (A1, A4) showed loss at the class-phylum level, indicating the presence of one or more phyla with more than one class. In general, a 'loss of information α' is evident in the 'inner group' if the benthic assemblages are studied at a higher taxonomic level than that of species. In contrast, in the 'outer group' most of the stations (A5, A6, A7, A8, A9, K3) show a 'loss of information α' only in the passage class-order level, because of the presence of only a single species for each genus; similarly, all families contained a single genus and all orders contained only a single family, displaying a value equal to zero for most of the levels considered. In this case there is a null 'loss of information α' if the benthic assemblage is studied at higher taxonomic level than species level.

Abundance matrices were analyzed from species to phylum level and six different ordination plots are shown (Fig. 2). Two different groups of stations (30% of similarity), corresponding to the 'inner' and 'outer' locations, are distinguished in the analysis performed from species to family level. Station A2 is an exception, given that it is located near the installation (50 m) but linked to the 'outer group'. Station A9 detaches from the 'outer group' at both species and genus levels, underlining its differentiation (Fig. 2). At higher taxonomic levels (order, class, phylum) there is reduction of information and it is no longer possible to identify the two clusters at the same percentage of similarity (Fig. 2). However, the 'inner group' and the 'outer group' remain distant, and the dissimilarity among clusters is significant ($P < 0.01$).

In case study A there is therefore a good correlation between species level and ordination plots at lower taxonomic levels (genus, family; $P > 0.94$), whereas at higher levels (order, class, phylum) the information about the structure of the benthic assemblages decreases ($P < 0.77$).
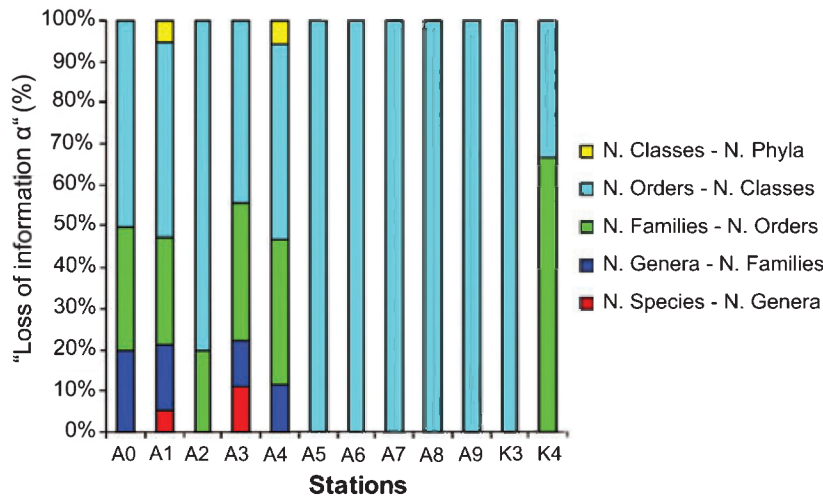
Fig. 1. 'Loss of information α' in percentage scale from lower to higher taxonomic levels at the 12 stations placed around the platform in case study A.
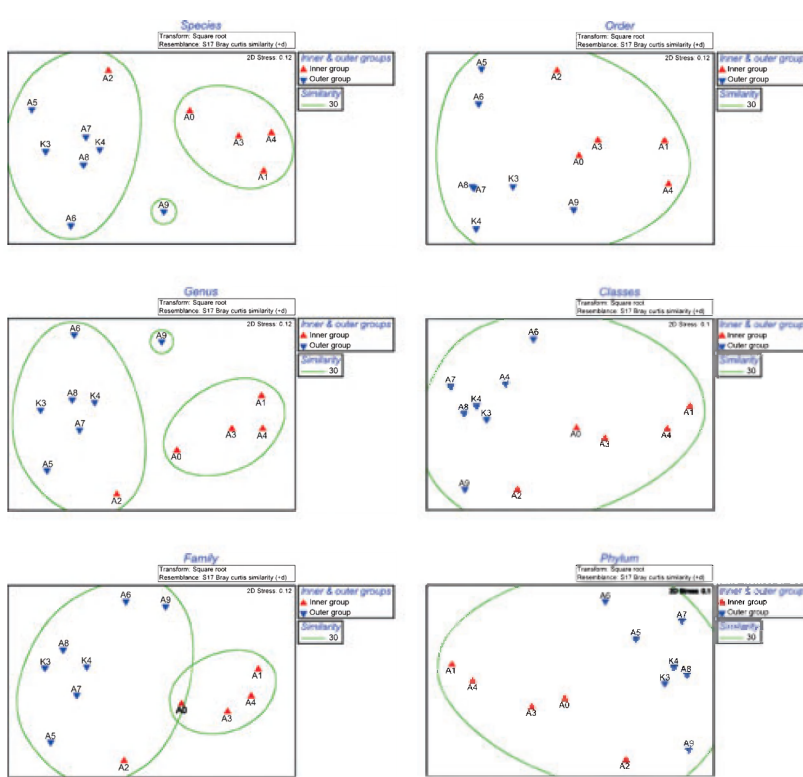

Fig. 2. nMDS and hierarchical cluster analysis at species, genus, family, order, class and phylum levels for case study A.

Case study B

Dataset B included a total of 12,318 individuals from the phyla Polychaeta, Mollusca, Crustacea and Echinodermata, and 225 taxa were identified at species level.

The study showed the presence of a classic soft-bottom benthic assemblage along a coast-wide transect sited in the North Adriatic Sea. Most of the species collected near the coast were characteristic of the Biocoenose des Sables Fins des Hauts Niveaux, SFHN, and the Biocoenose des Sables Fins Bien Calibrés, SFBC. At increasing depth there

was a group of stations where the presence of several exclusive/characteristic species of different biocoenoses (SFBC, VTC, DC) was recorded. This settlement is defined here as 'transitional stations' (TS). Most of the species collected far from the coast were characteristic of the Biocoenose de la Vase Terrigène Cotière, VTC, and the Biocoenose des Fonds Détritiques Cotière, DC (Vatova 1949; Peres & Picard 1964; Gamulin-Brida 1967; Meinesz et al. 1983; Bacci et al. 2009).

A high number of species and individuals and a complex taxonomic structure were recorded. It is possible to
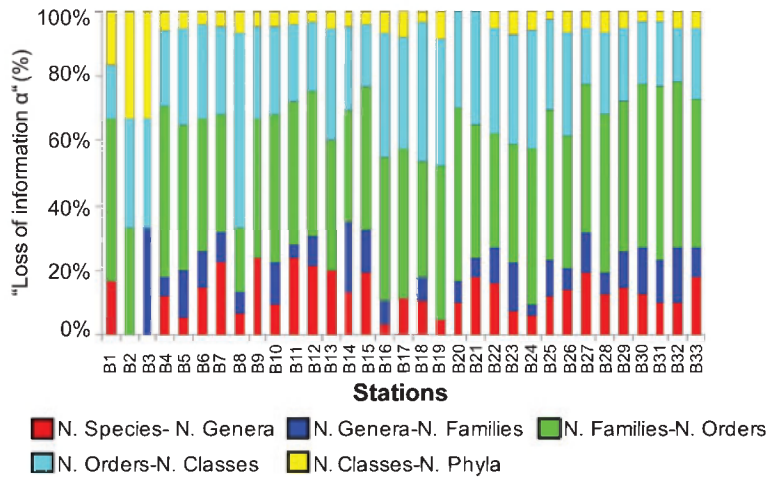
Fig. 3. 'Loss of information α' in percentage scale from lower to higher taxonomic level at the 33 stations along the coast-wide transect in case study B.

note up to five different species per genus and seven genera per family. This complex taxonomic structure occurred in most of the stations along the transect off-coast as highlighted by 'α' values (Fig. 3). 'Loss of information α' can be recorded in the passage between all levels considered in the majority of the stations (B4, B5, B6, B7, B8, B10, B11, B12, B14, B15, B16, B18, B22, B23, B24, B25, B26, B27, B28, B29, B30, B31, B32, B33). In fact, most of the benthic assemblages along the entire coast-wide transect showed the presence of one or more taxa within all taxonomic levels, from genus to phylum.

Abundance matrices have been analyzed from species to phylum level and six different ordination plots are shown (Fig. 4). Different groups of stations (40% of similarity) stand out in the analysis performed from species to family level (Fig. 4). At higher taxonomic levels there is a reduction of information and it is no longer possible to identify the stations near the coast (SFHN and SFBC)
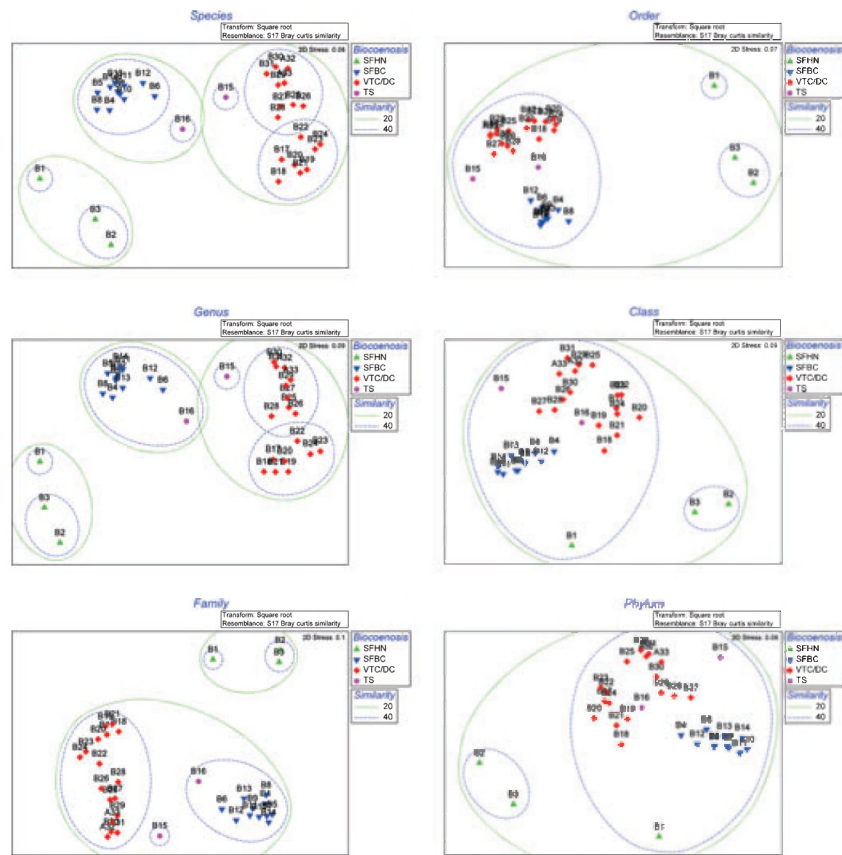

Fig. 4. nMDS and hierarchical cluster analysis at species, genus, family, order, class and phylum levels for case study B.

and the stations far off the coast (VTC, DC) at the same percent of similarity (Fig. 4). However, different biocoenoses remain distant, and the dissimilarity among clusters is significant (P < 0.01). It is interesting to observe that stations B1, B2 and B3 are grouped together from species to family level at 20% similarity because they actually belong to the same biocoenosis (Fig. 4). Nevertheless, B1 is separate from stations B2 and B3 at a higher similarity (40%); this is due to a quantitative difference caused by the great abundance of some species in this site. A second observation should be made of the VTC/DC stations. At species and genus levels the stations identified as VTC/DC group together with a 20% similarity. The same does not happen at family level, where they group together with a 40% similarity. This difference of information is due to the fact that at lower taxonomic levels the graphs highlight different proportions of species characteristic of VTC and of DC (more abundant in the stations more distant from the coast), which are instead disguised at family level.

In case study B there is a good correlation between species level and ordination plots at lower taxonomic levels (genus, family; P > 0.94), whereas at higher levels (order, class, phylum) the information about the structure of benthic assemblages decreases (P < 0.87).

Discussion

The two datasets differ for environmental conditions such as grain-size sediment, bathymetric gradient and human pressure. These diverse conditions contribute to the differentiation of the two benthic assemblages. The benthic assemblage along the off-coast transect appears to have a more complex structure than the benthic assemblage around the off-shore gas platform, due to a considerably higher number of species and individuals and therefore a more complex taxonomic structure.

Taxonomic heterogeneity is well described by 'α' values, which appear higher in case study B (i.e. higher loss of information on the taxonomic structure of benthic settlement) because of the general presence of more than one taxon within all taxonomic levels, from genus to phylum.

Nonetheless, taxonomic sufficiency seems to work in both sites, from species to higher taxonomic levels, as shown by ANOSIM. Other statistical analyses such as MDS, hierarchical clustering and Spearman's correlation suggest the family taxonomic level as the best compromise for taxonomic resolution (i.e. null reduction of information) when an accurate identification is not possible. MDS results display stress (S) values minus or near to 0.1 in both our cases, justifying a good measure of the degree of coupling of real between-sample distances and between-sample distances in the ordination

(Kruskal 1964; Gomez Gesteira et al. 2003) for each taxonomic level considered.

Contrary to expectations, we observed that in case study B there was not a high loss of information about the general structure of the benthic assemblages from species to family taxonomic level, despite the high number of species per genus and high number of genera per family. We could formulate the hypothesis that a high number of species, genera and families makes the benthic assemblage structure stable and diverse up to family level, compensating the loss of information at the passage to the higher taxonomic level.

The application of taxonomic sufficiency does not indicate a significant difference between the two datasets and appears therefore to be a valid instrument to analyze and describe the structure of benthic settlements in the case of taxonomically complex communities as well.

References

Anderson M.J., Connell S.D., Gillanders B.M., Diebel C.E., Bloms W.M., Saunders J.E., Landers T.J. (2005) Relationships between taxonomic resolution and spatial scales of multivariate variation. *Journal of Animal Ecology*, **74**, 636–646.

Bacci T., Trabucco B., Marusso V., Vani D., Lomiri S., Gabellini M., Virno Lamberti C. (2009) Caratterizzazione biocenotica di un'area dell'Alto Adriatico: risultati preliminari di uno studio sulla comunità macrozoobentonica marina di fondo mobile. Atti del congresso congiunto AIOL – SItE – Ancona 2007.

Bray J.R., Curtis J.T. (1957) An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs*, **27**, 325–349.

Carriker M. (1976) The crucial role of systematics in assessing pollution effects on the biological utilization of estuaries. *Estuarine pollution control assessment. Proceedings of Conference*, 487–506. USEPA Office of Water Planning and Standards.

Castanedo N.D., Lopez R.R., Weiss V.S., Alcantara P.H., Barba A.G. (2007) The use of higher taxa to assess the benthic conditions in southern Gulf of Mexico. *Marine Ecology*, **28** (Suppl. 1), 161–168.

Castelli A., Lardicci C., Tagliapietra D. (2003) Il macrobenthos di fondo molle. In: Gambi M.C., Dappiano M. (eds.) *Manuale di metodologie di campionamento e studio del benthos marino mediterraneo. Biologia Marina Mediterranea*, **10** (Suppl), 109–144.

Clarke K.R., Gorley R.N. (2006) *Primer v6. User Manual/Tutorial*. PRIMER-E, Plymouth: 190 pp.

Clarke K.R., Green R.H. (1988) Statistical design and analysis for a 'biological effects' study. *Marine Ecology Progress Series*, **46**, 213–226.

Clarke K.R., Somerfield P.J., Chapman M.G. (2006) On resemblance measures for ecological studies, including taxonomic

dissimilarities and zero-adjusted Bray-Curtis coefficient for denuded assemblages. *Journal of Experimental Marine Biology and Ecology*, **330**, 55–80.

Dauvin J.C., Gomez Gesteria J.L., Salvande Fraga M. (2003) Taxonomic sufficiency: an overview of its use in the monitoring of sublittoral benthic communities after oil spill. *Marine Pollution Bulletin*, **46**, 552–555.

Ellis D. (1985) Taxonomic sufficiency in pollution assessment. *Marine Pollution Bulletin*, **16**, 442–461.

Ferraro S.P., Cole F.A. (1992) Taxonomic level sufficient for assessing a moderate impact on macrobenthic communities in puget sound, Washington, USA. *Marine Ecology Progress Series*, **67**, 251–262.

Gamulin-Brida H. (1967) The benthic fauna of the Adriatic Sea. *Oceanography and Marine Biology. An Annual Review*, **5**, 535–568.

Gomez Gesteira J.L., Dauvin J.C., Salvande Fraga M. (2003) Taxonomic level for assessing oil spill effects on soft-bottom sublittoral benthic communities. *Marine Pollution Bulletin*, **46**, 562–572.

Gray J.S. (1981) *The Ecology of Marine Sediments*. Cambridge University Press, Cambridge: 185 pp.

Gray J.S. (1997) Marine biodiversity: patterns, threats and conservation needs. *Biological Conservation*, **6**, 153–165.

Gray J.S., Clarke K.R., Warwick R.M., Hobbs G. (1990) Detection of initial effects of pollution on marine benthos: an example from the Ekofisk and Eldfisk oilfields, North Sea. *Marine Ecology Progress Series*, **66**, 285–299.

Kruskal J.B. (1964) Non-metric multidimensional scaling: a numerical method. *Psychometrika*, **29**, 115–129.

Kruskal J.B., Wish M. (1978) *Multidimensional Scaling*. Sage, Beverly Hills, CA.

Meinesz A., Boudouresque C.F., Falconetti C., Astier J.M., Bay D., Blanc J.J., Bourcier M., Cinelli F., Cirik S., Cristiani G., Di Geronimo I., Giaccone G., Harmelin J.G., Laubier L., Lovric A.Z., Molinier R., Soyer J., Vamvakas C. (1983) Normalisation des symboles pour la représentation et la cartographie des biocénoses benthiques littorales de Méditerranée. *Annales de l'Institut Oceanographique*, **59**(2), 155–172.

Olsgard F., Somerfield P.J., Carr M.R. (1997) Relationships between taxonomic resolution and data transformations in analyses of macrobenthic community along an established pollution gradient. *Marine Ecology Progress Series*, **149**, 173–181.

Pearson T.H., Rosenberg R. (1978) Macrobenthic succession in relation to organic enrichment and pollution in the marine environment. *Oceanography and Marine Biology. An Annual Review*, **16**, 229–311.

Peres J.M., Picard J. (1964) Nouveau manuelle de bionomie bentique de la Mer Mediteranee. *Recueil des Travaux de la Station Marine d'Endoume*, **31**(47), 5–137.

Raymond W.B. Jr, Huggins D., Kriz J. (2005) A review of the issues related to taxonomic resolution in biological monitoring of aquatic ecosystem with an emphasis on macroinvertebrates. Prepared in fulfillment of USEPA Grant X7-99790401. Central Plains Center for BioAssessment, Kansas Biological Survey, Lawrence, KS.

Snelgrove P.V.R. (1997) The importance of marine sediment biodiversity in ecosystem processes. *Ambio*, **26**, 578–583.

Terlizzi A., Bevilacqua S., Fraschetti S., Boero F. (2003) Taxonomic sufficiency and the increasing insufficiency of taxonomic expertise. *Marine Pollution Bulletin*, **46**, 556–561.

Trabucco B., Cicero A.M., Gabellini M., Virno Lamberti C., Di Mento R., Bacci T., Moltedo G., Tomassetti P., Panfili M., Marusso V., Cornello M. (2006) Studio del popolamento macrozoobentonico di fondo mobile in prossimità di una piattaforma off-shore (Adriatico Centrale). *Biologia Marina Mediterranea*, **13**(1), 659–662.

Vanderklift M.A., Ward T.J., Jacoby C.A. (1996) Effect of reducing taxonomic resolution on ordinations to detect pollution-induced gradients in macrobenthic infaunal assemblages. *Marine Ecology Progress Series*, **136**, 137–145.

Vatova A. (1949) La fauna bentica dell'alto e medio Adriatico. *Nova Thalassia*, **1**(3), 110.

Warwick R.M. (1988) Analysis of community attributes of the macrobenthos of Frierfjord/Langesundfjord at taxonomic levels higher than species. *Marine Ecology Progress Series*, **46**, 167–170.

Warwick R.M. (1993) Environmental impact studies on marine communites: pragmatical considerations. *Australian Journal of Ecology*, **18**, 63–80.

Warwick R.M., Clarke K.R. (1991) A comparison of some methods for analysing changes in benthic community structure. *Journal of the Marine Biological Association of the United Kingdom*, **71**, 225–244.