

Relationships between taxonomic resolution, macrobenthic community patterns and disturbance

Frode Olsgard^{1,*}, Paul J. Somerfield², Martin R. Carr²

¹Section of Marine Zoology and Marine Chemistry, Department of Biology, University of Oslo, PO Box 1064, Blindern, N-0316 Oslo, Norway

²Plymouth Marine Laboratory, Prospect Place, The Hoe, Plymouth PL1 3DH, United Kingdom

ABSTRACT: Although analyses of macrobenthic infaunal communities traditionally require identification of the organisms to the species level, there is still only a limited number of studies which have examined the use of higher-level taxa in monitoring surveys. These generally indicate that identification of organisms to the lowest possible taxon may not always be necessary to enable description of spatial patterns in routine environmental and pollution monitoring programs. In other fields of marine ecology, such as studies of biodiversity, comparative investigations over large geographical areas, and the development of rapid-assessment techniques, the use of identification to higher taxonomic levels is often necessary. Detailed comparative studies of faunal patterns at various taxonomic levels are therefore important. This study uses multivariate analyses of macrofauna and environmental data from 20 separate investigations in the Skagerrak and North Sea to examine faunal patterns at different taxonomic levels. Data are analysed at the levels of species, genus, family, order, class and phylum. Transformation of data is also considered since degree of transformation is as important in determining the outcome of subsequent analyses as the taxonomic level to which organisms are identified. Correlations between the underlying similarity matrices at the species level and higher taxonomic levels show highest values in polluted areas, lower values in less disturbed areas and lowest values in pristine areas, indicating that the faunal patterns for the various taxonomic levels become more similar as the degree of disturbance increases. For those areas where time-series data are available it is shown that, as contamination increases, correlations between the faunal patterns at all taxonomic levels, but especially the highest levels, and the environmental variables tend to increase, and after several years of contamination all the correlations are high, independent of taxonomic level. Again, the faunal patterns at the different taxonomic levels tend to become more similar as a result of increased pollution. For all 20 investigations, highest correlations between faunal patterns (i.e. the underlying similarity matrices) and environmental variables were mainly found at the levels of species, genus and family, and often there was a distinct drop in correlation value between family and order. In most cases, independent of the level of pollution, there is only a minor reduction in correlation between species and family, suggesting that identification to the level of family may be satisfactory in many routine monitoring surveys. The debate about the level of taxonomic resolution required for routine environmental monitoring becomes relatively unimportant if the effects of choosing different transformations are not also considered.

KEY WORDS: Taxonomic resolution · Transformations · Community patterns · Macrobenthos · Multivariate analysis · Disturbance · Pollution monitoring · Cost-effectiveness

INTRODUCTION

Environmental monitoring and pollution assessment programmes, the main objective of which is often to identify the principal patterns in community structure

and to relate them to measured environmental variables, traditionally require that organisms be identified to the species level. There is still only a limited number of studies in which the use of abundances of higher-level taxa (such as genera and families) has been examined as an alternative to species-level identification. The potential usefulness of higher-level taxa

*E-mail: frode.olsgard@bio.uio.no

in these investigations is related to whether data at higher taxonomic levels still enable us to assess most of the variation in community structure. The important question, therefore, is to what taxonomic level is it necessary to identify organisms in order to meet, adequately, the objectives of a study. If appropriate multivariate analyses of data from a taxonomic level higher than species result in similar groupings, this would be a strong indication that higher-level taxa are sufficient in that particular investigation. Several authors have pointed out the need for further investigations of 'taxonomic sufficiency' (sensu Ellis 1985) before general recommendations are given for future macrobenthic monitoring surveys (e.g. Warwick 1988a, Gray et al. 1990, Vanderklift et al. 1996, Olsford et al. 1997).

Warwick (1988a) suggested that the results of multivariate analyses based on higher taxa may more closely reflect gradients of contamination than those based on species abundances, the latter being more affected by natural environmental variables. Several authors (e.g. Pearson & Rosenberg 1978, Boesch & Rosenberg 1981, Ferraro & Cole 1990) have suggested that the relevance of higher taxonomic levels may be a consequence of the hierarchical structure of biological responses to stress. In a hierarchic-response-to-stress hypothesis they propose that, as toxic stress increases, the adaptability of first the individual, and then species, genus, family etc., is exceeded, and therefore impacts resulting from increasing stress are manifest at higher and higher taxonomic levels. This implies that if disturbance increases in a particular area, the faunal gradients become stronger, and therefore identification of organisms to higher taxonomic levels should enable clearer taxonomic identification of changes in community structure.

Studies of patterns in biodiversity have received increasing interest during recent years, as the rapid loss of marine habitats, and thereby marine biodiversity, especially in coastal areas, has become of concern (Gray 1997), raising the importance of methods by which biodiversity may be investigated over larger geographical areas. Macrobenthic investigations using abundances of organisms at levels higher than species may seem contradictory in such studies, where identification to the species level is the norm. However, in most marine areas with high biodiversity, such as the tropics, there is also a lack of taxonomic literature and expertise. Furthermore, it is difficult to examine such patterns at the level of species where different areas have few or no species in common. To overcome these problems, such investigations can only be realised using identification to higher taxonomic levels. Comparative studies between regions using higher taxonomic levels include investigations from terrestrial (e.g. Prance 1994, Williams & Gaston 1994, Balmford et al. 1996), freshwater (e.g. Gaston et al. 1995, Williams

et al. 1997) and marine (e.g. Roy et al. 1996, Myers 1997) environments. Other marine examples include the investigation of longitudinal and latitudinal diversity gradients, described among others in Stehli et al. (1969), Clarke (1992), Etter & Grassle (1992) and Gray (1997). Also, in experimental work, higher taxonomic levels have been used as proxies for species-level identification (Morrisey et al. 1995, 1996).

Because attempts to catalogue the distribution of biological diversity may be facilitated by focusing at the level of genera or families rather than species, the development of rapid-assessment techniques, which also use identification of organisms to levels higher than species, will probably be an important, integral part of such studies. We will probably see, in environmental monitoring, biodiversity and rapid-assessment studies, an increased number of investigations performed using identification to higher taxonomic levels.

Although there is considerable debate about the validity of using taxonomic levels other than species in community investigations, there has been little consideration of other decisions made during the analyses of the data collected in such studies. For example, transformations are routinely employed in multivariate analyses to adjust the influences of common and rare taxa. Results of analyses using untransformed abundances reflect changes among the numerically dominant taxa, whereas a strong transformation, such as reducing abundances to presence/absence, gives all taxa an equivalent weighting regardless of their relative abundances. The choice of transformation used to analyse a set of data can influence the results of consequent analyses just as much as the choice of taxonomic level (Olsford et al. 1997), and it is important that investigations into the utility of higher-level taxa in monitoring surveys also consider relevant aspects of consequent analyses. It is time that more detailed, comparative studies of faunal patterns at the various taxonomic levels are undertaken to examine general patterns and enable recommendations for future work. The present study is a contribution to these investigations.

MATERIALS AND METHODS

Macrofaunal abundances (pooled data from either 4 or 5 replicate grabs from each station) from 20 separate investigations at 10 oil and gas fields in the North Sea and 2 areas in the Oslofjord, Norway, collected between 1985 and 1993 at depths between 14 and 380 m, are used in this study (Fig. 1, Table 1). A comprehensive analysis of most of these data was performed by Olsford & Gray (1995) in order to detect effects of pollution from offshore oil and gas activities on the macrobenthos.

Species abundances from each investigation were aggregated to the levels of genus, family, order, class and phylum following the taxonomic classification of Howson (1987). Subsequent calculations were done using the PRIMER package (Clarke & Warwick 1994). From the resulting 6 different abundance matrices for each investigation, a total of 24 ranked matrices of similarities among samples from each investigation were constructed using the Bray-Curtis similarity measure (Bray & Curtis 1957) and the following transformations: none, square root, fourth root and presence/absence. Multivariate patterns in these matrices were compared using the method of Somerfield & Clarke (1995), whereby rank correlations, calculated between pairs of similarity matrices, become the elements of a second similarity matrix, which is then used as an input matrix for a 'second-stage' MDS (multidimensional scaling) ordination. In such an ordination each symbol represents a similarity matrix, and the relative distances between symbols indicate how similar they are. Thus the patterns in a number of similarity matrices from each investigation can be compared simultaneously, e.g. to study the relative effects of aggregation and transformation on multivariate patterns.

The rank correlation used, prior to ordination, was a modified Spearman rank correlation, the harmonic rank correlation of Clarke & Ainsworth (1993), which downweights the effects of larger rank dissimilarities. This minimises the effect of outliers, which can have a large influence on Spearman rank correlation calculated between similarity matrices. Correlations (Spearman rank) between similarity matrices for species abundances and matrices derived from abundances of higher taxonomic levels were also calculated for each field, using RELATE (Clarke & Warwick 1994). Ordination was by non-metric MDS, and the goodness-of-fit of the resulting 2-dimensional plots was measured using Kruskal's Stress Formula I (Kruskal & Wish 1978, Clarke & Green 1988).

The relationships between patterns in multivariate community structure and environmental variables in each survey were examined using the BIO-ENV procedure (Clarke & Ainsworth 1993). This calculates rank correlations between a similarity matrix derived from

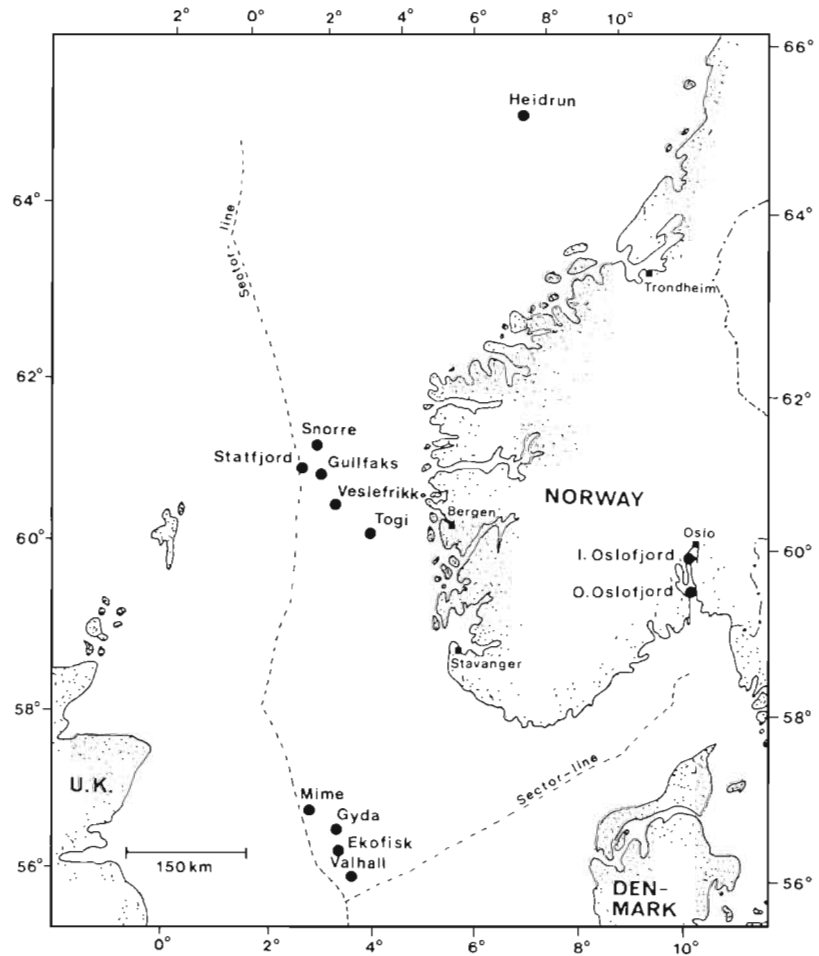


Fig. 1. Locations of the areas investigated (●) in the North Sea and Oslofjord (I: inner fjord; O: outer fjord)

biotic data and matrices derived from various subsets of environmental variables, thereby defining suites of variables most closely correlated with the observed biotic structure. The environmental variables were converted to approximate normality using a \log_{10} transformation prior to these analyses, and for each survey all 24 similarity matrices (6 taxonomic levels \times 4 transformations) were analysed separately. The actual environmental variables measured in each investigation, and the suite of variables most closely correlated with the biotic structure in each similarity matrix, are different. It is not the aim of the present paper to discuss the details of these relationships in any depth. Instead we make the assumption that, in all of the original investigations, a sensible and relevant set of environmental variables were chosen, and that by using BIO-ENV to select the most relevant subset of variables the goodness-of-fit between environmental variables and biotic structure is maximised. Links between faunal patterns and environmental variables

Table 1. Summary showing the 20 separate investigations and some basic data, with ranges where applicable. Total S: total number of species found at the area investigated; S: number of species per station; N: number of individuals per station (for Inner and Outer Oslofjord: per 0.4 m²); pollution history: number of years since contamination started; relative degree of pollution: 0 indicates an unpolluted area, 100 indicates a highly polluted area

Field/area	Year	No. of stations	Depth (m)	Total S	S (per 0.5 m ²)	N (per 0.5 m ²)	Pollution history (yr)	Degree of pollution
Heidrun	1988	25	320–380	192	46–81	141–448	0	0
Snorre	1989	40	288–340	319	59–114	162–525	0	0
Togi	1989	21	299–303	120	24–43	83–166	0	0
Mime	1990	11	80	162	51–99	319–528	0	0
Valhall	1985	13	65–68	73	4–34	74–502	3	28
Valhall	1988	20	65–69	96	21–47	201–1868	6	46
Valhall	1991	16	65–69	156	9–66	261–3583	9	58
Gyda	1987	18	65	106	26–43	102–209	0	0
Gyda	1990	17	65	139	14–70	136–1600	1	37
Gyda	1993	20	65	156	29–67	255–2270	4	28
Veslefrikk	1987	10	171–181	127	43–62	115–240	0	0
Veslefrikk	1990	16	171–186	238	71–104	231–3290	4	10
Veslefrikk	1993	14	171–186	267	61–122	507–11600	7	29
Gullfaks A	1989	16	131–140	249	42–109	264–1054	3	27
Gullfaks B	1993	13	133–196	248	58–116	441–2680	5	0
Ekofisk	1990	39	67–76	179	9–63	198–3400	20	34
Statfjord A	1993	11	140–144	225	50–101	464–9410	14	11
Statfjord C	1993	13	136–143	249	39–114	167–8570	8	56
O. Oslofjord	1988	19	24–355	200	32–78	231–2410	>10	10
I. Oslofjord	1993	42	14–203	257	1–92	2–4690	>20	50

are often used to find the variables 'best' explaining spatial or temporal patterns in benthic communities. The underlying assumption is that faunal patterns are a function of variation in environmental conditions. However, in the present paper we follow the method of Olsfjord et al. (1997), who used results from BIO-ENV to reveal which taxonomic level and transformation give the 'best' description of community patterns, defined as those giving the maximum correlation with the environmental variables.

RESULTS

A summary of basic data from the 20 surveys is shown in Table 1, in which pollution history indicates the number of years since contamination started. The surveys with 0 yr of pollution history are baseline studies performed before production of oil and gas commenced. Olsfjord & Gray (1995) used the number of species 500 m from the source of contamination as a percentage of the number of species at the control and/or unaffected station(s) to indicate the degree of pollution. We use a modified measure to indicate the degree of pollution (Table 1), in which a value of 0 indicates an unpolluted area, and 100 indicates a highly polluted area. The original investigations can be seen to span a range of pollution histories, from unpolluted areas to areas with a long history and severe degree of pollution.

One major consequence of identification to higher taxonomic levels is that abundances are discriminated into a smaller number of larger groups. If an effect of pollution is the alteration of the balance between numbers of species and numbers in higher taxonomic groups, this will have consequences for the way in which perceived patterns differ in analyses based on abundances of different taxonomic levels. As an example, if pollution reduces the number of species in each family to a single species, then in polluted areas analyses based on family abundances will be identical to those based on species abundances. The number of species, genera, families, orders, classes and phyla within each survey is given in Table 2. The number of families as a percentage of the number of species varies from 35.0% (Inner Oslofjord 1993) to 72.6% (Valhall 1985), with a mean value of 54.5% for all surveys, and is related to the total number of species within each survey (Pearson's correlation, $r = 0.90$, $n = 20$, $p < 0.001$). For the various surveys there is, however, no significant relationship between the number of families as a percentage of the number of species and the degree of pollution, indicated in Table 1 ($r = 0.16$, $n = 20$, $p = 0.515$).

In order to examine whether there are consistent differences in the way in which patterns change with changes in taxonomic level and transformation we use 'second-stage' MDS ordinations (Figs. 2 & 3). Each plot shows the interrelationships between the 23 of the 24

Table 2. Summary of areas of investigation, showing year of sampling and number of species, genera, families, orders, classes, and phyla within each area. Number of families as % of number of species in parentheses

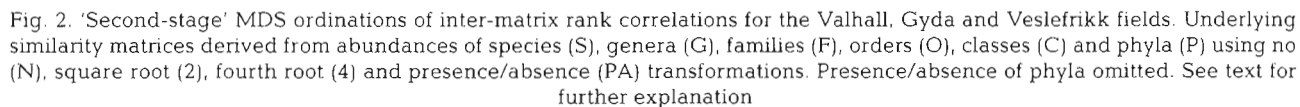
Field/area	Year	Species	Genera	Families (% of species)	Orders	Classes	Phyla
Heidrun	1988	192	156	111 (57.8)	54	19	10
Snorre	1989	319	238	135 (42.3)	51	17	11
Togi	1989	120	103	81 (67.5)	40	14	9
Mime	1990	162	131	94 (58.0)	37	16	10
Valhall	1985	73	67	53 (72.6)	28	13	9
Valhall	1988	96	85	65 (67.7)	31	15	9
Valhall	1991	156	138	102 (65.4)	42	18	10
Gyda	1987	106	90	64 (60.4)	28	16	9
Gyda	1990	139	116	88 (63.3)	33	15	9
Gyda	1993	156	128	92 (59.0)	39	18	10
Veslefrikk	1987	127	114	80 (63.0)	46	19	9
Veslefrikk	1990	238	179	104 (43.7)	50	20	9
Veslefrikk	1993	267	214	104 (39.0)	52	21	12
Gullfaks A	1989	249	198	125 (50.2)	61	26	13
Gullfaks B	1993	248	204	120 (48.4)	58	21	11
Ekofisk	1990	179	134	95 (53.1)	33	17	11
Statfjord A	1993	225	170	100 (44.4)	49	19	9
Statfjord C	1993	249	190	117 (47.0)	60	23	10
O. Oslofjord	1988	200	177	104 (52.0)	52	20	11
I. Oslofjord	1993	257	164	90 (35.0)	36	13	8

similarity matrices derived from the abundance data gathered in each study. The matrix derived from presence/absence of phyla is omitted in each case, as it is so different that its inclusion obscures the relationships between the others. For every survey the ordinations show 'fan patterns', with a horizontal separation of taxonomic levels and a vertical separation of transformations. This clearly shows that the effects of aggregation and transformation are to a large extent unrelated, confirming in more general terms the findings of Olsgard et al. (1997). Although there are differences in the shapes of the fan patterns, these do not seem to be linked to whether an area is clean or contaminated. In all plots it can be seen that matrices derived from species, genus and family abundances constructed using the same transformation tend to cluster together, indicating that the similarity matrices are closely related to each other. Transformation has more of an effect than the level to which organisms are identified (symbols are further apart), and as the taxonomic level increases the effects of subsequent transformations become stronger.

There is a relationship between the level of pollution in an area and correlations between similarity matrices derived from species abundances and those derived from abundances of higher taxa (Fig. 4). In polluted areas, such as Gyda 1993, Valhall 1991, Ekofisk 1990, Statfjord C 1993 and Veslefrikk 1993, the correlations between the species levels and higher taxonomic levels are relatively high (low slope), indicating that in polluted areas, where the initial faunal patterns are

strong, these patterns are evident when data are aggregated to higher taxonomic levels, even to the levels of class or phylum. For the baseline studies and in mildly polluted areas, where patterns are weak to begin with, the correlations are lower.

The hierarchic-response-to-stress hypothesis suggests that, as toxic stress increases, the adaptability of first the individual, and then species, genus, family, etc., is exceeded, and as a consequence changes resulting from increasing stress are manifest at higher and higher taxonomic levels. Thus it might be expected that the taxonomic level most highly correlated with measured environmental variables, including contaminants, will increase with increasing levels of stress. The results from BIO-ENV analyses of data from 3 subsequent investigations at each of the Valhall, Gyda and Veslefrikk fields (Table 3) indicate that there is no simple relationship between the time for which a stress is applied, or the degree of pollution, and the correlation values. For the Valhall field the highest correlations occur around the levels of genus, family or order, independent of transformation. At the Gyda and Veslefrikk fields the highest correlations occur with a wider range of taxonomic levels, depending on which transformation is used, and there is a tendency for correlation values at all taxonomic levels to increase as the degree of pollution increases through the years and as gradients are established. This implies that the faunal patterns at different taxonomic levels tend to become more similar. At the Valhall field these trends are less obvious, probably because high levels



To examine further whether a relationship exists between the taxonomic level most highly correlated with environmental variables and the level of pollu-

tion, additional data sets were investigated using BIO-ENV (Table 4). A square root transformation of faunal data, which is likely to give a balance between a 'narrow view' of community structure, highly influenced by dominant taxa, and a 'wide view' of community structure, in which all taxa contribute equally regard-

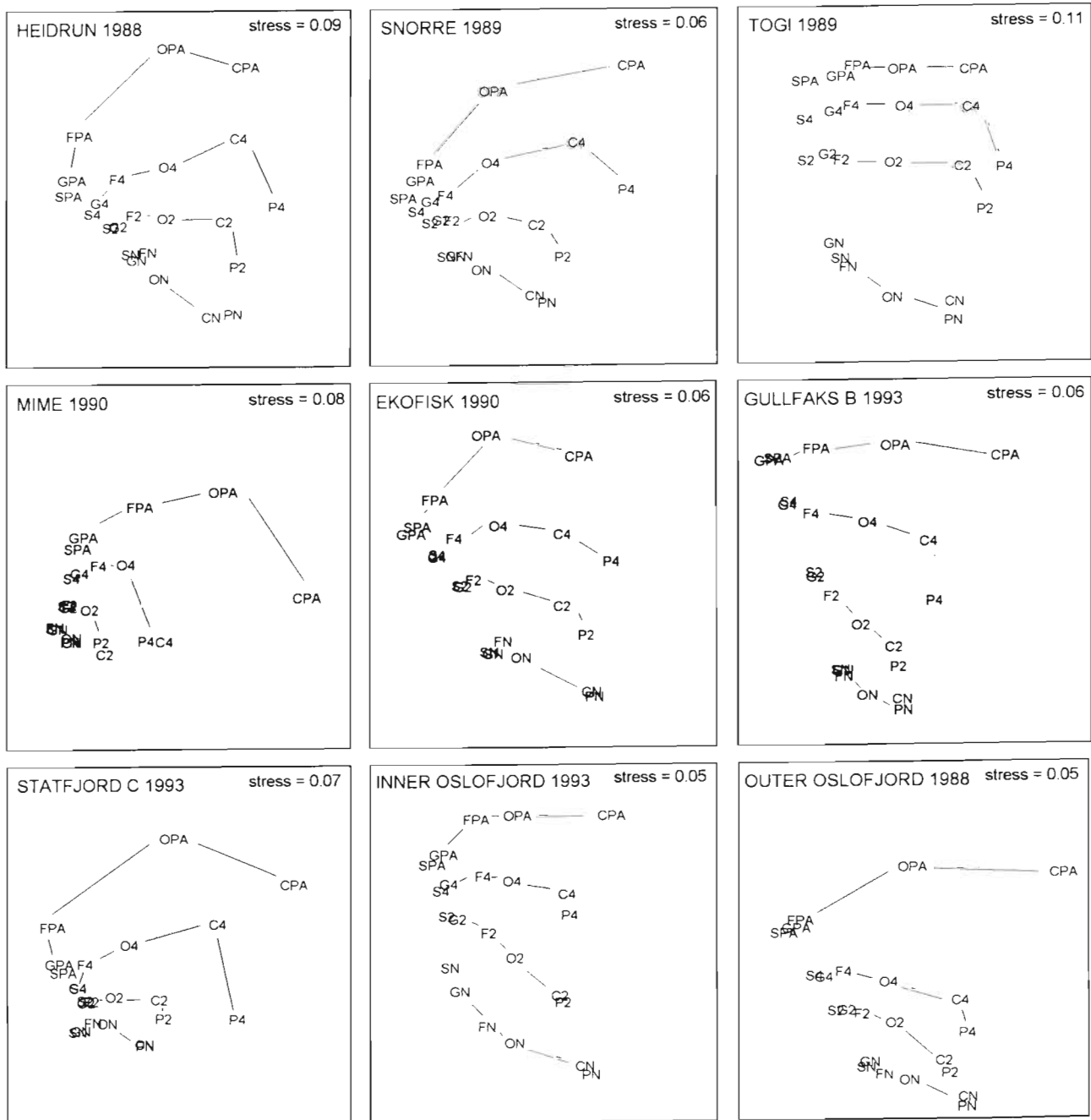


Fig. 3. 'Second-stage' MDS ordinations of inter-matrix rank correlations for the Heidrun, Snorre, Togi, Mime, Ekofisk, Gullfaks B, Statfjord C fields, and Inner and Outer Oslofjord. Abbreviations as in Fig. 2. Presence/absence of phyla omitted

less of their relative abundances (Clarke & Warwick 1994, Olsgard et al. 1997) was chosen for all investigations and taxonomic levels, while environmental data were \log_{10} transformed to obtain approximate normality. As expected, the lowest correlation values were found for baseline studies, e.g. Heidrun 1988, Mime 1990, Togi 1989, Snorre 1989, Gyda 1987, and Veslefrikk 1987. Higher correlations were found in areas

with clear pollution impacts, e.g. Valhall 1991, Gyda 1993, Veslefrikk 1993, Ekofisk 1990 and Statfjord C 1993. In areas where there is no clear relationship between the biotic and environmental variables, such as the baseline studies at the Heidrun, Mime, Togi and Snorre fields, no particular taxonomic level gives the maximum correlation consistently (Table 4). The low correlations indicate that in these areas biological

Table 3. Summary of results from BIO-ENV for data from the Valhall, Gyda and Veslefrikk fields comparing environmental variables with fauna to show the taxonomic level and transformation giving the 'best' description of community patterns in each field. Maximum correlation values for separate transformations in **bold**; maximum correlation values each year, independent of transformation, are underlined

Year	Transformation	Species	Genus	Family	Order	Class	Phylum
Valhall							
1985	None	0.727	0.717	0.797	0.692	0.473	0.475
	Square root	0.795	0.799	0.773	0.729	0.665	0.652
	Fourth root	0.801	0.825	0.724	0.724	0.663	0.690
	Presence/absence	0.783	0.795	0.704	0.704	0.648	0.656
1988	None	0.673	0.668	0.678	0.620	0.376	0.377
	Square root	0.693	0.703	0.711	0.737	0.465	0.429
	Fourth root	0.717	0.721	0.714	0.681	0.469	0.352
	Presence/absence	0.609	0.628	0.597	0.569	0.447	0.142
1991	None	0.797	0.800	0.801	0.774	0.605	0.605
	Square root	0.833	0.849	0.850	0.847	0.752	0.736
	Fourth root	0.798	0.814	0.810	0.827	0.799	0.745
	Presence/absence	0.713	0.725	0.728	0.658	0.699	0.412
Gyda							
1987	None	0.421	0.425	0.441	0.429	0.373	0.388
	Square root	0.312	0.318	0.322	0.336	0.344	0.388
	Fourth root	0.259	0.255	0.274	0.342	0.379	0.321
	Presence/absence	0.242	0.229	0.259	0.348	0.357	0.398
1990	None	0.718	0.722	0.702	0.712	0.612	0.608
	Square root	0.743	0.743	0.722	0.742	0.719	0.662
	Fourth root	0.721	0.712	0.708	0.751	0.775	0.661
	Presence/absence	0.628	0.573	0.634	0.705	0.738	0.664
1993	None	0.765	0.772	0.776	0.748	0.737	0.724
	Square root	0.773	0.752	0.784	0.718	0.720	0.737
	Fourth root	0.745	0.718	0.741	0.635	0.668	0.674
	Presence/absence	0.654	0.605	0.608	0.413	0.283	0.256
Veslefrikk							
1987	None	0.371	0.371	0.401	0.384	0.429	0.411
	Square root	0.398	0.403	0.522	0.377	0.418	0.584
	Fourth root	0.476	0.423	0.555	0.325	0.273	0.539
	Presence/absence	0.464	0.432	0.536	0.216	0.199	0.443
1990	None	0.762	0.761	0.754	0.730	0.382	0.411
	Square root	0.733	0.729	0.714	0.658	0.521	0.565
	Fourth root	0.698	0.674	0.616	0.669	0.518	0.626
	Presence/absence	0.638	0.591	0.413	0.229	0.220	0.254
1993	None	0.676	0.667	0.662	0.648	0.661	0.667
	Square root	0.765	0.751	0.718	0.653	0.636	0.636
	Fourth root	0.815	0.804	0.754	0.638	0.605	0.616
	Presence/absence	0.852	0.825	0.747	0.525	0.504	0.467

interactions are more important structuring factors than variation in environmental variables. Strong correlations, including those with higher taxonomic levels, are generally found in areas with a longer pollution history, e.g. Valhall 1991, Gyda 1993, Veslefrikk 1993, Ekofisk 1990 and Statfjord C 1993. In most cases the maximum correlation values were found with the species level, but it should be noted that there is only a minor drop in correlation values between the levels of species, genus and family. Thus the identification of organisms to the levels of genus or family, or even to higher taxonomic levels, still results in interpretable

faunal ordinations in polluted areas. For the majority of the surveys examined the most obvious drop in correlation values occurs between the taxonomic levels of order and class, which is in agreement with the pattern evident in Fig. 4.

DISCUSSION

In analyses of data from the 1991 survey of the Valhall field, aggregated to various taxonomic levels, Ols-gard et al. (1997) showed that choices of both taxo-

nomic level and transformation were important in affecting the results of subsequent analyses. The effects of each were different and, to a large extent, unrelated as 'second-stage' MDS of inter-matrix rank correlations showed orthogonal separation of groups by taxonomic levels and by transformations. In the present study, results from both offshore and inshore areas showed similar 'fan patterns' to that found for Valhall 1991 (Figs. 2 & 3). No evidence was found to suggest an impact of pollution on the general pattern. It is shown that for each separate transformation the ordinations derived from abundances of species, genera and families are, in most cases, similar (symbols close together). At the level of order the differences from these lower taxonomic levels become more obvious (symbols further apart), and these differences increase as abundances are aggregated to classes and phyla.

Warwick (1988a, b) and Gray et al. (1990), using visual examinations of

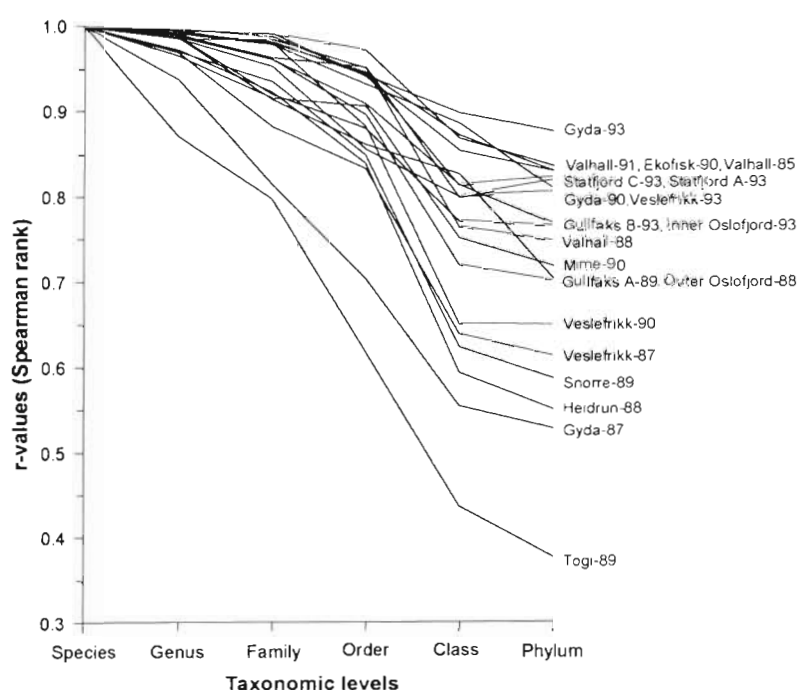


Fig. 4. Degree of correlation between the underlying faunal similarity matrices for the species level and the matrices for the higher taxonomic levels at each field. Data are square root transformed

Table 4. Results from BIO-ENV, comparing environmental variables with fauna for the 20 investigations, to show the taxonomic level and transformation giving the 'best' description of community patterns in each area. Maximum correlation values for each investigation in **bold**. Faunal data are square root transformed. Degree of pollution as in Table 1

Field/area	No. of stations	Species	Genus	Family	Order	Class	Phylum	Degree of pollution
Heidrun 1988	25	0.477	0.478	0.459	0.435	0.336	0.277	0
Mime 1990	11	0.351	0.305	0.322	0.360	0.365	0.456	0
Togi 1989	21	-0.023	-0.012	0.048	0.124	0.157	0.053	0
Snorre 1989	40	0.452	0.469	0.427	0.350	0.188	0.165	0
Valhall 1985	13	0.795	0.799	0.773	0.729	0.665	0.652	28
Valhall 1988	20	0.693	0.703	0.711	0.737	0.465	0.449	46
Valhall 1991	16	0.833	0.849	0.850	0.847	0.752	0.736	58
Gyda 1987	18	0.312	0.318	0.322	0.336	0.334	0.388	0
Gyda 1990	17	0.743	0.743	0.722	0.742	0.719	0.662	37
Gyda 1993	20	0.773	0.752	0.784	0.718	0.720	0.737	28
Veslefrikk 1987	10	0.398	0.403	0.522	0.377	0.418	0.411	0
Veslefrikk 1990	16	0.733	0.729	0.714	0.658	0.521	0.565	10
Veslefrikk 1993	14	0.765	0.751	0.718	0.653	0.636	0.636	29
Gullfaks A 1989	16	0.572	0.572	0.550	0.504	0.525	0.423	27
Gullfaks B 1993	13	0.676	0.657	0.614	0.544	0.506	0.453	0
Ekofisk 1990	39	0.809	0.808	0.773	0.737	0.644	0.634	34
Statfjord A 1993	11	0.846	0.863	0.844	0.772	0.769	0.750	11
Statfjord C 1993	13	0.821	0.822	0.820	0.821	0.630	0.675	56
O. Oslofjord 1988	19	0.735	0.718	0.676	0.601	0.434	0.394	10
I. Oslofjord 1993	42	0.603	0.597	0.539	0.510	0.446	0.423	50

ordination plots, concluded that analyses of data at the levels of class and phylum were able to reveal the main faunal patterns in the areas of investigation. Using more objective analyses in investigations of macrobenthos from the clearly polluted Liverpool Bay and Fal estuary, UK, Somerfield & Clarke (1995) found that correlations between matrices derived from abundances of species and abundances of phyla were $r = 0.750$ and $r = 0.799$, respectively, and they concluded that results from analyses conducted at the highest taxonomic levels could be interpreted. In the survey of Valhall in 1991 (Olsgard et al. 1997) correlations between the underlying similarity matrices derived from species abundances and those from abundances of genera, families and orders were high ($r = 0.99$, 0.99 and 0.97 , respectively) but lower for abundances of classes and phyla ($r = 0.91$ and 0.89 , respectively). They also concluded that, given an established pollution gradient in an area, the main faunal patterns can still be discovered using abundances of higher taxonomic levels, even at levels as high as class or phylum. The present study seems to confirm the generality of these findings. In polluted areas, such as Gyda 1993, Ekofisk 1990, Statfjord C 1993 and Veslefrikk 1993, high correlations between species and higher taxonomic levels are found. In less perturbed, but still polluted areas with clear gradients in the community structure of the benthic fauna, such as Gullfaks B 1993, Inner Oslofjord 1993 and Valhall 1988, the correlations between the species level and the levels of class and phylum are somewhat lower, compared to the most affected areas, implying that lower taxonomic levels should be used. In baseline studies, such as Snorre 1989, Heidrun 1988, Gyda 1987 and Togi 1989, there was an obvious drop in correlation values between matrices derived from species abundances and those derived from all higher taxonomic levels. For these surveys, the levels of class and phylum correlations were always lower than $r = 0.70$, implying that in areas with weak gradients in faunal composition, analyses at the levels of class or phylum should be avoided. In general there was a gradual increase in the correlation values between abundances at the species level and the various higher taxonomic levels for the fields where pollution impact increased over time (Valhall, Gyda and Veslefrikk). This clearly lends support to the hierarchical-response-to-stress hypothesis (Pearson & Rosenberg 1978, Boesch & Rosenberg 1981, Ferraro & Cole 1990).

Although identifying organisms to a taxonomic level higher than species, to family for example, potentially shortens the time needed to identify the animals and thereby reduces the cost of a survey, this will only be of use if the assumption underlying most environmental surveys, namely that faunal patterns are a function

of variations in environmental conditions and that changes in the environment elicit detectable responses in the fauna, holds true for abundances of higher taxa. Olsgard et al. (1997) used the correlation values from BIO-ENV to examine which combination of taxonomic level and transformation gave the best match with the measured environmental variables for Valhall 1991, as a method to reveal the 'best' taxonomic level and transformation to choose to describe the community patterns within a particular area, and found that the maximum correlation was for square root transformed abundances at the family level. Table 3 clearly indicates that the choices of both taxonomic level and transformation have impacts on the correlation values, but in general as the degree of pollution increases the correlation values for all taxonomic levels tend to increase, suggesting that where gradients are established investigations based on higher taxonomic levels are likely to be of use in describing the main patterns in faunal composition.

Warwick (1988a, b, 1993) suggested that community responses to pollution should be more evident at higher taxonomic levels than species since analyses based on species abundances are more likely to be complicated by natural environmental heterogeneity than analyses of higher taxonomic levels. Thus community responses to human perturbations may be detected more easily above the noise of natural variability by working at high taxonomic levels, especially in geographically large and environmentally heterogeneous areas. This idea is supported by Smith & Simpson (1993) in a study of effects of pollution on kelp holdfast macrofauna in New South Wales, Australia, although Wright et al. (1995) came to the opposite conclusion in a study of benthic macroinvertebrates in upland streams in New South Wales. In the present study we investigated data from surveys representing disturbance gradients from geographically limited areas, which were perhaps too small to test this idea properly. We note, however, that for those investigations from environmentally heterogeneous areas such as Veslefrikk, Gullfaks and Statfjord the species or genus levels give the highest BIO-ENV correlation values (Table 4), and the level of phylum gives the lowest correlations. Thus these analyses do not support the idea that higher taxonomic levels in heterogeneous environments reflect the effects attributable to perturbation more clearly, although they may do so over larger areas.

In all the 20 investigations examined in this study, independent of levels of pollution, both 'second-stage' MDS (Figs. 2 & 3) and BIO-ENV (Tables 3 & 4) show that aggregation of data to the level of family in most, not to say all, cases produces results which are very close to those based on species-level identification, adding weight to the findings of previous studies that

data aggregated to family level are no less informative than species-level data in studies aimed at detecting community patterns (Warwick 1988a, b, Ferraro & Cole 1990, 1995, James et al. 1995, Somerfield & Clarke 1995, Vanderklift et al. 1996, Olsgard et al. 1997). Although these findings could be used to make a general recommendation that the family level is a suitable level for faunal identification in studies of marine macrofauna, this will to a large extent depend on the objectives of each individual investigation. In baseline studies, and ecologically orientated surveys, the identification of organisms to the level of species is highly recommended. Only with such data is a complete knowledge of the pre-impact community composition possible, revealing to what extent there are species of particular ecological importance (keystone species) or potential commercial value in the area in question. In biodiversity studies, where differing areas being compared have few or no species in common, analyses based on higher taxonomic levels are necessary. Most examinations of taxonomic levels in macrobenthic community studies have, as here, been performed on data collected in monitoring surveys in areas affected by pollution. A limited identification of organisms, to the family level for example, is of particular interest in repeated routine investigations. In benthic sampling programmes a great portion of the resources is spent in identifying organisms in the laboratory and there is a high potential for time and cost savings through the use of coarser taxonomic resolution. Time and money saved could also be allocated elsewhere, for example to allow for more samples to be taken, giving better spatial coverage of the area under study. For the 20 investigations in the present study, aggregation of data from species to family level reduced the mean number of taxa to 54.5% (Table 2), which indicates a saving potential of almost 50% by using family instead of species identification. However, if unexpected trends in benthic community patterns are discovered using abundances of families, then identification can still be continued to the level of species in order to elucidate the details of the detected changes.

The finding of Olsgard et al. (1997) that the degree of transformation chosen is as important to the outcome of consequent analyses as the taxonomic level to which the fauna is identified is shown to be true for all the investigations examined in this study. Of course it is important to realise that abundance data are not transformed to satisfy statistical assumptions, but simply to adjust the contributions of taxa to inter-sample similarities (Clarke 1993). With no transformation, only a few dominant taxa will contribute to the similarities, whereas with a presence-absence transformation all taxa are equally important. The choice of transforma-

tion is, therefore, one that should be made with due consideration by the investigator.

In monitoring programmes it is often the case that all individuals are identified to species level and then the data are routinely analysed using a fourth root transformation. If no other transformation need be considered, then time can be saved during data collection by noting that a fourth root transformation is roughly equivalent to reducing abundances to a scale of 0 = absent, 1 = 1 individual, 2 = a few individuals, 3 = several individuals, 4 = abundant, ≥ 5 = very abundant (Clarke & Warwick 1994); hence, rather than counting all the individuals in samples it would only be necessary to assess their abundance on a crude scale and then use the (untransformed) values in subsequent multivariate analyses.

Acknowledgements. As a result of the efforts of various consultants, institutions and oil companies, many data have emerged from offshore biological and chemical monitoring of the Norwegian Continental shelf, and we gratefully acknowledge their work. We also thank 3 anonymous referees for their valuable criticism. Parts of this study were supported by grants to F.O. from the Norwegian Research Council and NATO. This study was also funded in part by the UK Ministry of Agriculture, Fisheries and Food (Project AE 1113) and is a contribution to PML's Coastal Biodiversity research project.

LITERATURE CITED

- Balmford A, Green MJB, Murray MG (1996) Using higher-taxon richness as a surrogate for species richness. 1. Regional tests. *Proc R Soc Lond B* 263:1267–1274
- Boesch DF, Rosenberg R (1981) Response to stress in marine benthic communities. In: Barrett GW, Rosenberg R (eds) *Stress effects on natural ecosystems*. Wiley, New York, p 179–200
- Bray JR, Curtis JT (1957) An ordination of the upland forest communities of southern Wisconsin. *Ecol Monogr* 27: 325–349
- Clarke A (1992) Is there a latitudinal diversity cline in the sea? *Trends Ecol Evolut* 7:286–287
- Clarke KR (1993) Non-parametric multivariate analyses of changes in community structure. *Aust J Ecol* 18:117–143
- Clarke KR, Ainsworth M (1993) A method for linking multivariate community structure to environmental variables. *Mar Ecol Prog Ser* 92:205–209
- Clarke KR, Green RH (1988) Statistical design and analysis for a 'biological effects' study. *Mar Ecol Prog Ser* 46: 213–226
- Clarke KR, Warwick RM (1994) *Changes in marine communities: an approach to statistical analysis and interpretation*. Plymouth Marine Laboratory, Plymouth
- Ellis D (1985) Taxonomic sufficiency in pollution assessment. *Mar Pollut Bull* 16:439
- Etter RJ, Grassle JF (1992) Patterns of species diversity in the deep sea as a function of sediment particle size diversity. *Nature* 360:576–578
- Ferraro SP, Cole FA (1990) Taxonomic level and sample size sufficient for assessing pollution impacts on the Southern California Bight macrobenthos. *Mar Ecol Prog Ser* 67: 251–262

- Ferraro SP, Cole FA (1995) Taxonomic level sufficient for assessing pollution impacts on the Southern California Bight macrobenthos—revisited. *Environ Toxicol Chem* 14:1031–1040
- Gaston KJ, Williams PH, Eggleton P, Humphries CJ (1995) Large scale patterns of biodiversity: spatial variation in family richness. *Proc R Soc Lond B* 260:149–154
- Gray JS (1997) Marine biodiversity: patterns, threats and conservation needs. *Biodivers Conserv* 6:153–175
- Gray JS, Clarke KR, Warwick RM, Hobbs G (1990) Detection of initial effects of pollution on marine benthos: an example from the Ekofisk and Eldfisk oilfields, North Sea. *Mar Ecol Prog Ser* 66:285–299
- Howson CM (1987) Species directory to British marine fauna and flora. Marine Conservation Society, Ross-on-Wye
- James RJ, Lincoln Smith MP, Fairweather PG (1995) Sieve mesh-size and taxonomic resolution needed to describe natural spatial variation of marine macrofauna. *Mar Ecol Prog Ser* 118:187–198
- Kruskal JB, Wish M (1978) Multidimensional scaling. Sage Publishers, Beverly Hills, CA
- Morrisey DJ, Underwood AJ, Howitt L (1995) Development of sediment-quality criteria: a proposal from experimental field-studies of the effects of copper on benthic organisms. *Mar Pollut Bull* 31:372–377
- Morrisey DJ, Underwood AJ, Howitt L (1996) Effects of copper on the faunas of marine soft sediments: an experimental field study. *Mar Biol* 125:199–213
- Myers AA (1997) Biogeographic barriers and the development of marine biodiversity. *Estuar Coast Shelf Sci* 44:241–248
- Olsgard F, Gray JS (1995) A comprehensive analysis of the effects of offshore oil and gas exploration and production on the benthic communities of the Norwegian continental shelf. *Mar Ecol Prog Ser* 122:277–306
- Olsgard F, Somerfield PJ, Carr MR (1997) Relationships between taxonomic resolution and data transformations in analyses of a macrobenthic community along an established pollution gradient. *Mar Ecol Prog Ser* 149:173–181
- 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
- Prance GT (1994) A comparison of the efficacy of higher taxa and species numbers in the assessment of biodiversity in the neotropics. *Proc R Soc Lond B* 345:89–99
- Roy K, Jablonski D, Valentine JW (1996) Higher taxa in biodiversity studies: patterns from eastern Pacific marine molluscs. *Proc R Soc Lond B* 351:1605–1613
- Smith SDA, Simpson RJ (1993) Effects of pollution on hold-fast macrofauna of the kelp *Ecklonia radiata*: discrimination at different taxonomic levels. *Mar Ecol Prog Ser* 96:199–208
- Somerfield PJ, Clarke KR (1995) Taxonomic levels, in marine community studies, revisited. *Mar Ecol Prog Ser* 127:113–119
- Stehli FG, Douglas RG, Newell ND (1969) Generation and maintenance of gradients in taxonomic diversity. *Science* 164:947–949
- Vanderklift MA, Ward TJ, Jacoby CA (1996) Effect of reducing taxonomic resolution on ordinations to detect pollution-induced gradients in macrobenthic infaunal assemblages. *Mar Ecol Prog Ser* 136:137–145
- Warwick RM (1988a) Analysis of community attributes of the macrobenthos of Frierfjord/Langesundfjord at taxonomic levels higher than species. *Mar Ecol Prog Ser* 46:167–170
- Warwick RM (1988b) The level of taxonomic discrimination required to detect pollution effects on marine benthic communities. *Mar Pollut Bull* 6:259–268
- Warwick RM (1993) Environmental impact studies on marine communities: pragmatical considerations. *Aust J Ecol* 18:63–80
- Williams PH, Gaston KJ (1994) Measuring more of biodiversity: can higher-taxon richness predict wholesale species richness? *Biol Conserv* 67:211–217
- Williams PH, Gaston KJ, Humphries CJ (1997) Mapping biodiversity value worldwide: combining higher-taxon richness from different groups. *Proc R Soc Lond B* 264:141–148
- Wright IA, Chessman BC, Fairweather PG, Benson LJ (1995) Measuring the impact of sewage effluent on the macroinvertebrate community of an upland stream: the effect of different levels of taxonomic resolution and quantification. *Aust J Ecol* 20:142–149

Editorial responsibility: Otto Kinne (Editor),
Oldendorf/Luhe, Germany

Submitted: March 16, 1998; Accepted: July 20, 1998
Proofs received from author(s): September 28, 1998