OPEN ACCESS

Macroecology of the European soft sediment benthos: insights from the MacroBen database

**T. J. Webb[1],*, I. F. Aleffi, J. M. Amouroux, G. Bachelet, S. Degraer, C. Dounas,
D. Fleischer, A. Grémare, M. Herrmann, H. Hummel, I. Karakassis, M. Kędra,
M. A. Kendall, L. Kotwicki, C. Labrune, E. L. Nevrova, A. Occhipinti-Ambrogi,
A. Petrov, N. K. Revkov, R. Sardá, N. Simboura, J. Speybroeck, G. Van Hoey,
M. Vincx, P. Whomersley, W. Willems, M. Włodarska-Kowalczuk**

[1]Department of Animal & Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK

ABSTRACT: Macroecology provides a novel conceptual framework for analysis of the distribution and abundance of organisms at very large scales. Its rapid development in recent years has been driven primarily by studies of terrestrial taxa; the vast potential of marine systems to contribute to the macroecological research effort remains largely untapped. International collaborative efforts such as MarBEF have provided fresh impetus to the collation of regional databases of species occurrences, such as the newly available MacroBen database of the European soft sediment benthic fauna. Here, we provide a first macroecological summary of this unique database. We show that in common with almost all previously analysed assemblages, the frequency distribution of regional site occupancies across species in the MacroBen database is strongly right-skewed. More unusually, this right skew remains under logarithmic transformation. There is little evidence for any major differences between higher taxa in this frequency distribution (based on the 8 animal classes for which we have sufficient data). Indeed, considerable variation in occupancy persisted across the taxonomic hierarchy, such that most variation occurred between species within genera. There was a weak positive relationship between local population density and regional occupancy across species, but this abundance–occupancy relationship varied considerably between higher taxa and between geographical areas. Our results highlight the potential of databases such as MacroBen to consolidate macroecological generalities and to test emerging theory.

KEY WORDS: Marine macroecology · Macrobenthic · Europe · Large marine ecosystems · Species–range size distributions · Phylogeny · Abundance–occupancy relationships

INTRODUCTION

Macroecology draws on insights from fields including ecology, biogeography, palaeontology, macroevolution and applied statistics to understand how large-scale processes affect the organisation of ecological systems at multiple scales (Brown 1995, Gaston & Blackburn 2000, Blackburn & Gaston 2006). It has defined novel and important concepts and methodological techniques to describe the form and structure of large-scale ecological patterns and has developed in a relatively short time into a thriving and productive discipline (Gaston &

Blackburn 2000, Blackburn & Gaston 2003). The importance of a macroecological approach becomes still more apparent with the realisation that human impacts on ecological systems are detectable at the same very large scales that interest macroecologists (Chapin et al. 2000, Kerr et al. 2007), and that many of the most pressing issues in applied ecology involve very general questions relating to habitat modification, invasive species, over-exploitation, pollution and climate change (Sutherland et al. 2006). In a marine context, considerable large-scale changes in ecosystems have already occurred without rigorous documentation (Jackson

2001), which brings into sharp relief the need to understand large-scale patterns of biodiversity that we may better predict the consequences of current and future human-induced global change.

A guiding principle of the macroecological approach is the search for generality — to what extent are large-scale patterns in the abundance and distribution of organisms consistent across taxa and environments? Certain macroecological patterns and relationships appear to be very general. For instance, across a taxonomically constrained group of species, the frequency distribution of range sizes measured at regional to global scales will typically display a characteristic right skew: most species are rare, but a few are extremely widely distributed (Gaston 2003). Similar generalities are observed in relationships between geographic distribution and local population density (Gaston et al. 2000, Blackburn et al. 2006). This so-called abundance–occupancy relationship is typically positive such that locally abundant species tend to be geographically widespread (Gaston et al. 2000, Blackburn et al. 2006), and is important as it provides a link between local and regional population processes (Freckleton et al. 2005, 2006).

An important caveat to the inferred generality of such patterns, however, is that macroecology has developed as an overwhelmingly terrestrial discipline (Raffaelli et al. 2005, Clarke et al. 2007), whereas the sea is home to most of life's higher-taxon diversity (May 1994). Only a minority of studies making explicit reference to 'macroecology' have had a primarily marine focus (Raffaelli et al. 2005), and just 18 of the 279 abundance–occupancy relationships reviewed by Blackburn et al. (2006) derived from marine or intertidal systems. This paucity of marine macroecological studies can be partially explained by the particular practical, logistical and financial challenges posed by working in the marine environment that make the kinds of datasets relied upon by terrestrial macroecologists (often compiled largely through the efforts of committed volunteer naturalists) difficult to obtain for marine taxa. In addition, it can prove difficult to define variables crucial to macroecological analyses such as 'range size' and 'population size', particularly for highly mobile pelagic taxa or for regions with (often highly) incomplete sampling. Despite these difficulties, there exist several excellent macroecological studies of the marine environment (e.g. Li 2002, Foggo et al. 2003, Macpherson 2003, Fisher & Frank 2004, Irigoien et al. 2004, Jennings & Blanchard 2004, Hsieh et al. 2006). In the main, these have tended to support general conclusions reached by terrestrial macroecologists on the basic form of the principal patterns. For instance, species–range size distributions display a similar characteristic right skew for both marine and terrestrial taxa (Gaston 2003), and similar patterns are also seen for species–abundance distributions; Gray et al. (2006) show that these are of similar form in a variety of marine and terrestrial taxa under similar sampling regimes. Importantly, macroecological relationships such as the abundance–occupancy relationship are also observed in the sea (e.g. Foggo et al. 2003, Fisher & Frank 2004, Blackburn et al. 2006), indicating that similar processes may act on communities in both environments. Efforts to increase communication between marine and terrestrial ecological research communities (e.g. Stergiou & Browman 2005) are stimulating progress in marine macroecology, and this trend will only increase as national and international organisations consolidate the data collected by individual research organisations into regional databases that provide a broader-scale overview of the distribution of marine diversity. The Marine Biodiversity and Ecosystem Functioning EU Network of Excellence (MarBEF, www.marbef.org) has been established specifically to investigate patterns of marine biodiversity, with an emphasis on the analysis of large-scale patterns and processes that would not be tractable without such international cooperation.

It is important to realise that macroecology is about more than simply describing patterns (Blackburn & Gaston 2006), and recent efforts have shifted towards seeking a mechanistic understanding of the processes that drive these patterns. For instance, population models that view interspecific macroecological patterns as the result of intraspecific dynamics have provided a powerful approach for generating mechanistic hypotheses (He & Gaston 2003, Freckleton et al. 2005, 2006), and make testable predictions about the roles of ecological and life history parameters in shaping macroecological patterns and relationships. Testing these predictions in terrestrial systems has been limited by the fact that parameters thought to be important in determining the form of macroecological patterns, such as colonisation ability, often do not vary substantially within the highly mobile taxa (such as birds) typically studied (Freckleton et al. 2005). Indeed most macroecological datasets tend to be somewhat taxonomically restricted, and thus encompass only limited functional diversity. This underlines the vast potential of macroecological studies of marine systems to aid in a more general understanding of macroecological process. The same sampling regime within a single marine habitat type will typically capture significant taxonomic diversity, recording species that differ markedly in characteristics (e.g. larval dispersal potential) hypothesised to be important in determining large-scale patterns of distribution and abundance (Foggo et al. 2007). Similarly, samples will often traverse steep environmental gradients (e.g. depth, salinity, fishing intensity). Because experimental manipula-

tions at macroecological scales are largely precluded, testing macroecological hypotheses across taxa and along environmental gradients may provide the most powerful test of their generality; this principle is taken to its extreme by testing in marine systems hypotheses generated from studies of terrestrial taxa (Steele 1991).

A necessary prerequisite to exploiting this vast potential of marine databases for testing emerging macroecological theory is a rigorous description of the observed patterns. Here, we analyse the MacroBen database to provide the first such summary for the soft-sediment macrobenthos throughout European coastal waters. We start by defining a simple measure of regional distribution based on presence/absence across a 100 km grid, and consider the frequency distributions of occupancy across 2292 species. We then use the exceptional taxonomic coverage of the MacroBen database to compare occupancy patterns among higher taxa. Specifically, we compare occupancies between 8 classes of animals for which we have distribution records for sufficient species, and provide the first analysis with this taxonomic scope of the way that variance in range size is partitioned across the taxonomic hierarchy. Finally, we introduce measures of population density allowing us to quantify abundance–occupancy relationships across and within classes, accounting for differences in mean density between geographic areas.

DATA AND METHODS

The MacroBen database. Producing comprehensive regional-scale databases through the integration of smaller-scale datasets has been a core objective of the MarBEF network. The MacroBen database is the first such regional database, comprising distribution records for soft sediment benthic taxa throughout European coastal waters compiled from many sources and published accounts including Karakassis & Eleftheriou (1997) and Mackie et al. (1995). The systematic nomenclature of the database has been checked for both orthography and synonymy against the European Register of Marine Species (ERMS, www.marbef.org/data/erms. php). Raw data were retained whenever possible to allow maximum flexibility during analyses, and geographical and physical data were available for large numbers of records. The final database contains 465 354 distribution records, from 7203 valid taxa and 22 897 sampling stations, collated from 44 individual data sets. It is presented as an MS Access file, and is described in full in Vanden Berghe et al. (2009, this Theme Section).

The MacroBen database includes tools to extract data and to calculate basic statistics and diversity coefficients. We applied the following filtering rules: only taxa identified to the species level were considered,

and samples were required to contain quantitative data on species numbers so that analyses using abundance data could proceed on the same set of species as analyses of occupancy patterns. Taxa not considered to be part of the macrobenthos were excluded, as were immature individuals. Only records collected since 1990 were retained, to reduce the influence of temporal trends in species distributions. All records were required to contain information on the area sampled, to allow estimates of population density (ind. m^{-2}) to be calculated for each species. All included datasets sampled using Van Veen grabs or hand-operated Van Veen grabs, generally with a sampling area of 0.1 m^2 (range 0.04 to 1 m^2, including pooled samples, and the mesh size used to sort samples was always 1 mm. Other potentially confounding factors remain in the dataset, either because their variation over large scales is seen as important in structuring macroecological relationships (e.g. spatial variability in habitat type) or because we had insufficient information to control for them (e.g. variability in the seasonality of sample collection). Our filtering rules resulted in a database containing 211 518 records, representing 2292 species from 15 024 sampling stations distributed between 6.5° W and 58.9° E and between 35.3° N and 81.5° N.

Measures of occupancy and population density. Site occupancy ('area of occupancy' in the terminology of Gaston 2003) is regularly used in macroecology as a measure of the extent of the distribution of a species. The simplest measure of site occupancy in our dataset would be the number (or proportion) of sampling stations at which a species was recorded. However, sampling effort was highly heterogeneous across space. To address this issue, we superimposed a 100 × 100 km grid onto the entire sampling area, projected using the Lambert Azimuthal Equal Area projection. Each individual spatially-referenced record could then be assigned to a single 100 × 100 km grid square, and occupancy was measured as the proportion of the total number of grid squares covered by the sampling area (n = 99) in which a given species was recorded. There are problems with this approach, particularly as some grid squares have been more heavily sampled than others. Thus, while presences can be interpreted with confidence, absences cannot; it is likely that most species occupy more squares than those in which they have been recorded. However, we assumed that the recorded occupancy of a species will be roughly in proportion to its actual occupancy. Thus, our occupancy estimates should give a reasonable approximation of the relative commonness or rarity of a species. The scale of grid chosen was also made primarily on pragmatic grounds, with 100 × 100 km resolution providing a good compromise between fine resolution and reasonable sampling effort. Fortunately, macroecological patterns

have proven to be remarkably robust to different scales of occupancy measure (Blackburn et al. 2004).

Population densities were calculated at the sampling station level, i.e. the total number of individuals of a species recorded across all replicate samples at a station was divided by the total area sampled. The mean density of a given species is taken as the geometric mean of its density estimates across all samples in which it was recorded (i.e. excluding zero densities).

A common feature of marine datasets is the presence of singletons, that is species recorded in the dataset as a single individual. Although they frequently form an important component of benthic samples, it has been argued that such very rare species may not play a significant role within communities, and that most are likely to be vagrants or transients that have immigrated from outside the sampling area (Gray et al. 2006). To quantify the effect of such species, we defined 2 kinds of singleton: (1) species represented by only a single individual across all replicates and sampling stations, and (2) species with a maximum abundance of 1 in any single sample. Whilst we recognise that sampling effects may create some 'artefactual' singletons (e.g. some large-bodied organisms may be inefficiently sampled, yet have important ecological interactions), using this compilation of multiple datasets means that the probability of erroneously excluding such rare or poorly-sampled important species is minimised.

Statistical analyses. The shape of the untransformed and $\log_{10}$-transformed frequency distributions of occupancies were quantified by their skewness, using the unbiased estimator of skewness (g_1) and its standard error given by Sokal & Rohlf (1995), which can be tested against a null expectation of 0 using the t distribution with df = ∞ (Sokal & Rohlf 1995). We performed this analysis for occupancies of all species, and for all non-singleton species (using the definitions of singletons described above).

Taxonomic patterns in occupancy were examined by first identifying those classes that contained at least 20 species (Table 1). These 8 classes, although representing just 24% of the classes present in the total database, together contained 93% of all recorded species and 97% of all distribution records. The variance in occupancy within this reduced dataset was partitioned between the levels of taxonomic hieararchy, i.e. class, order, family and genus, using a fully nested model, with variance components estimated using restricted maximum likelihood (REML). This procedure essentially partitions variance in occupancy into between and within group components at each subsequent taxonomic level. A high value for order, for example, would indicate that a large proportion of total variation in occupancy occurs between orders, with little variation within orders (i.e. species within an or-

Table 1. Number of species in each class in the complete dataset. The 8 classes shown in bold, with at least 20 species recorded in each, were used in analyses of individual classes

Phylum	Class	No. of species
Annelida	**Polychaeta**	**724**
Arthropoda	**Malacostraca**	**656**
Mollusca	**Bivalvia**	**290**
Mollusca	**Gastropoda**	**240**
Bryozoa	**Gymnolaemata**	**128**
Echinodermata	**Stelleroidea**	**50**
Echinodermata	**Holothuroidea**	**28**
Chordata	**Ascidiacea**	**27**
Cnidaria	Hexacorallia	19
Echinodermata	Echinoidea	17
Arthropoda	Pycnogonida	15
Bryozoa	Stenolaemata	14
Sipuncula	Sipunculidea	11
Mollusca	Aplacophora	10
Mollusca	Polyplacophora	9
Annelida	Clitellata	8
Cnidaria	Octocorallia	8
	Others (n = 14)	38
	Total	2292

der tend to be similar in terms of their occupancy). The residual variation from the model represents variation occurring between species within genera; a high value would indicate that even very closely related species (congeners) can differ substantially in their occupancies. This analysis is equivalent to the nested ANOVAs used for similar purposes elsewhere (e.g. Gaston 1998, Webb et al. 2001, Qian & Ricklefs 2004), although the REML estimation is more robust with unbalanced designs (Rao & Heckler 1997). Results need to be interpreted with caution, however, due to the influence of a large number of monotypic higher taxa: 19 of 69 orders (28%), 138 of 365 families (38%) and 584 of 1003 genera (58%) in the dataset were monotypic. Clearly, no variation is possible within a monotypic higher taxon, and such taxa will tend to inflate the proportion of variance explained at higher taxonomic levels. Our estimate of residual variation (i.e. the degree to which occupancy varies among congeneric species) is therefore bound to be an underestimate.

We estimated abundance–occupancy relationships using the Pearson product moment correlation between $\log_{10}$(occupancy) and $\log_{10}$(mean density). Although such relationships are frequently non-linear, the correlation coefficient provides a good approximation of their general form and strength (Webb et al. 2007). We first considered the relationship across all species in the dataset. We next examined the structure of this relationship by considering separately each of the 8 classes described above. In an attempt to control for differences in species composition and/or mean density across regions, we also estimated abundance–

occupancy relationships separately for each of the large marine ecosystems (LMEs, Sherman 1991) covered by our dataset. This involved calculating a separate occupancy and density value for each species in each LME. We excluded the Black Sea LME from this analysis, as all samples fell within a single 100 km square, meaning that there was no variance in occupancy within this LME. Finally, we combined the taxonomic and geographic stratification to estimate an abundance–occupancy relationship separately for each class within each LME.

All data manipulations and statistical analyses were executed with R 2.2.1 (R Development Core Team 2005, available at: www.R-project.org). The variance components analyses made use of the *nlme* package described in Pinheiro & Bates (2002).

RESULTS

Occupancy

The frequency distribution of occupancy was highly significantly right-skewed (skewness ± SE = 2.99 ± 0.051, t = 58.5, p < 0.00001; Fig. 1A). Substantial right-skew remained after log-transformation (skewness = 0.48 ± 0.051, t = 9.32, p < 0.00001; Fig. 1B). The vast majority of species, therefore, had extremely restricted recorded distributions; 680 species (30% of all species) occurred in only a single 100 km square. In contrast, the most widespread species (the polychaete *Heteromastus filiformis*) has been recorded in 73% of the sampled 100 km squares. The skew of the distributions does not result simply from the inclusion of 'accidental' species, at least if these were identified according to our definitions of singletons; significant right skew remained in untransformed and transformed distributions after the exclusion of both types of singleton (Fig. 1; excluding singletons with total count = 1, skewness of untransformed and log-transformed occupancy = 2.85 ± 0.054 and 0.35 ± 0.054; excluding singletons with maximum count = 1, skewness = 2.67 ± 0.058 and 0.22 ± 0.058; t > 3.7 and p < 0.0001 in all cases). Note that although 'true' singletons (those recorded as only a single individual across the entire dataset, n = 248 species) necessarily occur in only 1 square, species with a total abundance >1 but a maximum recorded abundance of 1 can be rather widespread (proportional occupancy ranged from 0.01 to 0.18; n = 253 species). We therefore consider these latter species to be integral, if under-sampled, components of the communities, and in subsequent analyses 'singleton' refers only to the former kind (only a single individual recorded).

Occupancy remained very variable, and highly skewed (skewness > 1.3, p < 0.0001) within each of the 8 classes that contained at least 20 species. After log transformation, all 8 distributions remained right-skewed (skewness ranged from 0.10 to 0.98), significantly so in the Ascidiacea, Gastropoda, Gymnolaemata and Malacostraca (Fig. 2A). The minimum
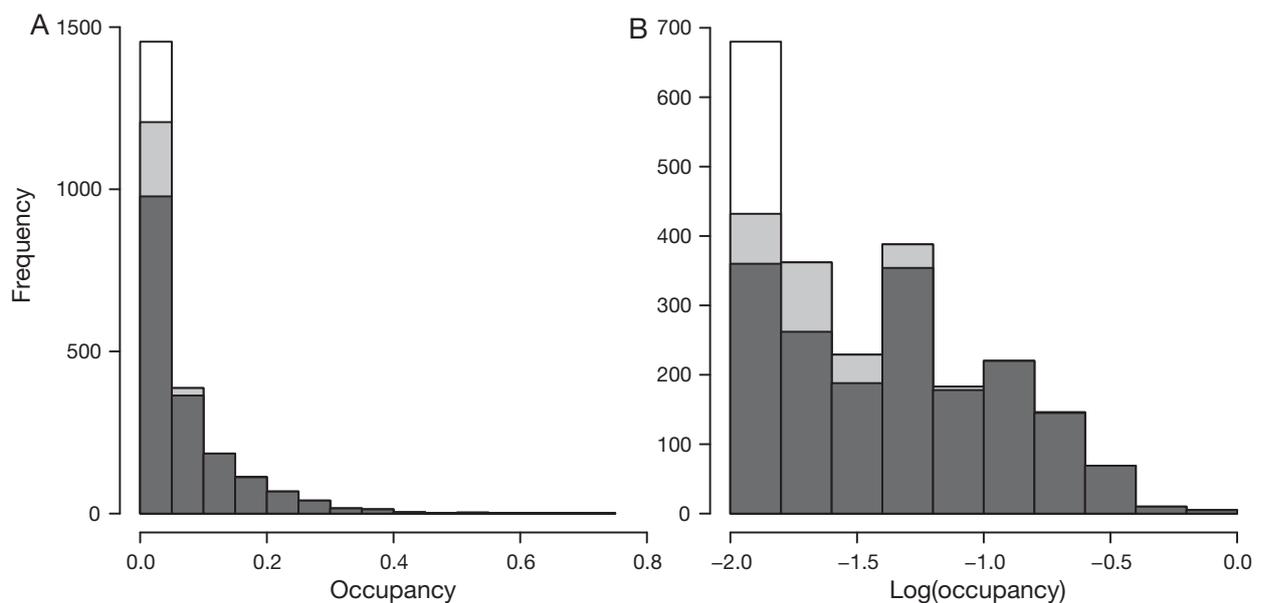

Fig. 1. Frequency distribution of occupancies (proportion of sampled 100 km squares in which a species was recorded) of European soft-sediment macrobenthic species. (A) occupancy untransformed; (B) occupancy log_{10}-transformed. In both cases, the total heights of bars represent the distributions for all 2292 species in the database; unshaded portions of the bars represent singletons, based on a total recorded count across all samples of 1 individual (n = 248); and light-shaded portions of the bars represent singletons that had a total abundance >1, but a maximum count in any one sample of 1 (n = 253)

observed occupancy was 0.01 (a single 100 km square) in all 8 classes. Polychaeta was the only class in which occupancies >0.35 occurred (in the total dataset, but not included in this class-level analysis, the sipunculid *Phascolion strombi* was the most widespread non-polychaete with an occupancy of 0.42). Maximum occupancies in 5 of the remaining 7 classes varied only between 0.26 and 0.35, with somewhat lower maximum occupancies in Class Ascidiacea (0.12) and Class Gymnolaemata (0.18). Overall, these results suggest that there was considerable variation in geographic distribution within classes. The taxonomically nested model shows that this variation persisted across the taxonomic hierarchy, such that the majority (73.3%) of variance in occupancy occurred between species within genera (Fig, 2B), despite the confounding issues surrounding the inclusion of monotypic higher taxa discussed in 'Data and methods'. Note that the proportion of variance explained at each taxonomic level was barely altered when singletons were excluded, and when monotypic genera were excluded (residual ['species within genera'] variation = 74.7 and 71.4%, respectively).

Abundance–occupancy relationships

Estimated population densities across the 2292 species in the total dataset spanned 6 orders of magnitude. The gastropods *Lacuna pallidula* and *Trophonopsis truncatus* and the polychaete *Orbinia norvegica* occurred at densities of just a single individual in 93 m², while the bivalve *Chamelea gallina* occurred at a geometric mean density of >1600 ind. m^{-2} across the 38 sampling stations at which it was recorded.

Although the correlation between mean population density and the proportion of 100 km squares occupied (across all species) was significantly positive (both variables $\log_{10}$-transformed, r = 0.22, df = 2290, p < 0.00001), the overriding impression obtained from Fig. 3 is of a relationship characterised by enormous variation. This impression is strengthened if singletons (which by definition have very low densities and can occupy only a single square) are excluded (r is reduced to 0.11 for the remaining 2044 species). In contrast to positive abundance–occupancy relationships observed in other systems, the pattern here was far from linear; low occupancies were observed across the range of densities, but high occupancies only occurred at intermediate to high densities.

One potential reason for the large amount of scatter in this cross-species relationship is that it is confounded by differences between higher taxa. As shown above, occupancy is not strongly constrained by taxonomy (closely related species can differ markedly in occupancy), but it is possible that different classes will
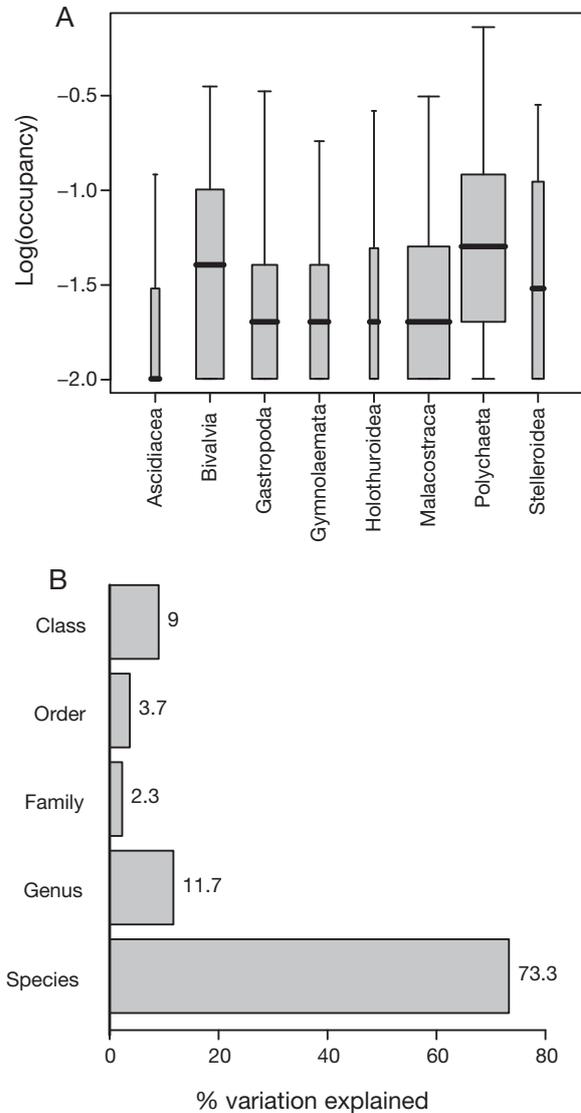

Fig. 2. (A) Box and whisker plots showing the distribution of $\log_{10}$-transformed occupancies within each of 8 classes in the database that contained >20 recorded species. The horizontal lines show the median occupancy for each class, boxes show the interquartile range, and whiskers extend to the data extremes. The width of each box is proportional to the number of species in that class. (B) Percentage of variation in $\log_{10}$-transformed occupancy explained at taxonomic levels from class to genus; unexplained (residual) variation is labelled 'Species', as it represents variation between species within genera. Variance components were estimated using restricted maximum likelihood, and the analysis included the 2143 species that occurred within the 8 classes illustrated in (A)

differ in the typical density achieved at a given level of occupancy (as predicted by population models for taxa with different population structures and life histories; Freckleton et al. 2005, 2006), thus introducing scatter into the cross-species relationship. In fact, mean population density varied by orders of magnitude between species within each class (Table 2), and in all classes
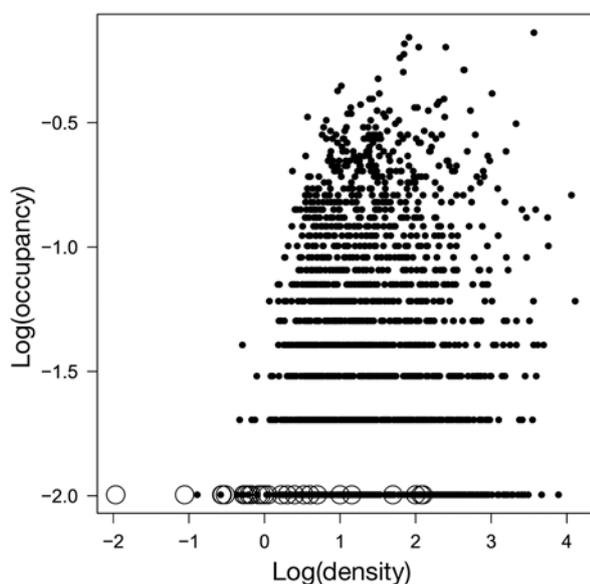
Fig. 3. Abundance–occupancy relationship for 2292 species in the database. The occupancy of a species is the proportion of 100 km squares in which it occurred; its density is the mean population density (ind. m^{-2}) in samples in which it was recorded. ○: singletons (species recorded as only a single individual across all samples; n = 248). Note that singletons can differ in density due to differences in the area sampled at different sampling stations

abundance–occupancy relationships were weak, with a maximum correlation of 0.28 in the Stelleroidea (Table 2). Although the within-class relationships were generally positive (in 5 of 8 classes, mean correlation [weighted by the square root of the number of species within each class] = 0.03), they typically retained the roughly triangular shape of the relationship across classes, with wide variation in occupancy across all except the lowest observed densities. Indeed, when singletons were removed, only 3 of 8 relationships remain positive, the strongest correlations were negative (Table 2), and the weighted mean correlation was negative (–0.09).

Another potential reason for the generally weak abundance–occupancy relationships concerns differences between LMEs in the typical densities attained by benthic species, which may be influenced by differences in, for example, productivity, habitat heterogeneity, ocean circulation, frequency or intensity of disturbance, and broad-scale heterogeneity in sampling effort. We therefore estimated separately for each LME the abundance–occupancy relationships across species. Here, oc-

cupancy was the proportion of 100 km squares within an LME occupied by a species, and density was its mean population density within those occupied squares. Again, mean density varied considerably between species within each LME (Table 3). Abundance–occupancy relationships were positive in all LMEs except the North Sea (Table 3), with a maximum correlation of 0.52 in the Baltic Sea and a weighted mean value of 0.21. Excluding singletons weakened all relationships (although that in the North Sea became more negative), and reduced the weighted mean correlation to 0.07, but it did not result in any changes of sign (Table 3). Again, however, all relationships contained more scatter than is frequently seen in other systems (e.g. Blackburn et al. 2006).

Finally, we investigated the interaction between taxonomy and geography by estimating abundance–occupancy relationships separately for each class in each LME. We only estimated relationships for classes with at least 8 species in any given LME. The majority of relationships (29 of 35) were positive, with a weighted mean correlation of 0.18, a pattern that held when singletons were removed (21 of 34 positive, weighted mean correlation = 0.04). The interaction between taxonomy and geography appears complex, however, as the rank order of classes in terms of their abundance–occupancy correlation differed between LMEs (Fig. 4). This suggests that the link between local and regional population processes in the marine benthos is influenced by both biotic and abiotic factors.

DISCUSSION

This first analysis of the macroecology of the European marine macrobenthic fauna has upheld several generalities familiar to macroecologists working in other systems. Thus, the frequency distribution of spe-

Table 2. Pearson product moment correlation coefficients for the relationship between log(population density) and log(occupancy) separately for each of 8 classes with >20 species, both including and excluding singletons. Also shown are the numbers of species in each class (number excluding singletons), and the range across species within each class in geometric mean population density (ind. m^{-2})

Class	No. of species	Abundance–occupancy correlation		Range in geometric mean density (ind. m^{-2})
		All species	Excluding singletons	
Polychaeta	724 (675)	0.13	0.02	0.01–1635
Malacostraca	656 (580)	–0.08	–0.19	0.10–1433
Bivalvia	290 (276)	–0.02	–0.08	0.53–1673
Gastropoda	240 (196)	–0.17	–0.36	0.01–873
Gymnolaemata	128 (102)	0.14	–0.13	0.28–1235
Stelleroidea	50 (43)	0.28	0.18	0.76–1151
Holothuroidea	28 (23)	0.22	0.31	0.78–21
Ascidiacea	27 (21)	0.05	–0.15	0.85–617

Table 3. Pearson product moment correlation coefficients for the relationship be-
tween log(population density) and log(occupancy) separately for each of the
6 large marine ecosystems (LMEs) for which relationships could be estimated. Cor-
relations were performed both including and excluding singletons. Also shown are
the number of species recorded in each LME (number excluding singletons), and
the range across species within each LME in geometric mean population density

LME	No. of species	Abundance– occupancy correlation		Range in geometric mean density (ind. m^{-2})
		All species	Excluding singletons	
Mediterranean Sea	947 (839)	0.24	0.15	0.09–860
Celtic-Biscay Shelf	494 (411)	0.22	0.09	3.13–3100
North Sea	1096 (954)	−0.08	−0.22	0.63–1660
Baltic Sea	222 (195)	0.52	0.41	0.01–1149
Norwegian Sea	492 (398)	0.26	0.13	0.91–73
Barents Sea	873 (728)	0.28	0.09	1.00–1187

cies occupancies displayed the strong right-skew typi-
cal of regional studies in a broad range of taxa in both
terrestrial and aquatic systems using various measures
of geographic distribution (e.g. Gaston 2003, Macpher-
son 2003, Clarke et al. 2007). Most species in the
MacroBen database were narrowly distributed (or at
least, have been recorded in only a few locations),
whereas some were much more widely distributed.
Importantly, even the most widely distributed species
(the polychaete *Heteromastus filiformis*) did not reach
full occupancy, suggesting that this analysis is truly
large-scale and captures a range of environmental
conditions broader than that which can be occupied by
most individual species. At smaller scales, species–
range size distributions can tend to bimodality, as sev-
eral widespread species occur in all sampled locations
(e.g. Storch & Sizling 2002).

Whilst the untransformed frequency distributions of
occupancies for European benthic taxa conformed to
macroecological expectation, the fact that they retained
right-skew under a logarithmic transformation is more
unusual. Typically, species–range size distributions ac-
quire a moderate left (negative) skew after such a trans-
formation (Gaston 2003, Macpherson 2003). A potential
explanation for this difference is that our occupancy
measures were at a finer scale (i.e. 100 km squares oc-
cupied, as opposed to degrees of latitude spanned), and
thus use more comprehensive distributional informa-
tion than previous regional-scale marine analyses (e.g.
Macpherson 2003). At the same time, the fauna in ques-
tion (European macrobenthic taxa) is considerably
richer and has been less exhaustively sampled than
many of the faunas considered in terrestrial macroeco-
logical analyses (e.g. birds and mammals). The expec-
tation would be that with continued sampling, the
right-skew of the distribution would decrease as cur-
rently 'rare' species were recorded at more locations (as

happened, for example, when we ex-
cluded singletons), although given that
additional sampling would likely also
uncover further singletons, the amount
of extra sampling required may prove to
be very large. We would expect that a
regional collation of incompletely sam-
pled, species-rich terrestrial datasets,
for instance tropical forest inverte-
brates, would display patterns similar to
those observed in the MacroBen data-
base (see Gray et al. 2006 for a discus-
sion of the influence of sampling
regime in marine and terrestrial sys-
tems on a related macroecological pat-
tern, the species–abundance distribu-
tion).

Patterns of regional occupancy in
European macrobenthic taxa do not show strong taxo-
nomic patterns; in all of the 8 classes we analysed,
occupancy was extremely variable (Fig. 2A). Impor-
tantly, this variation persisted down the taxonomic
hierarchy, with nearly 75% of interspecific variation in
occupancy occurring between species within genera
(Fig. 2B). Such patterns are typical of diverse assem-
blages for which such an analysis has been attempted
(e.g. Gaston 1998, Webb et al. 2001, Qian & Ricklefs
2004) and add further evidence that geographic distri-
bution is not highly constrained by phylogeny (Webb &
Gaston 2003, 2005, but see Hunt et al. 2005 for a
counter-argument).

Across all species in the MacroBen database, there
was a positive relationship between local abundance
(mean population density at occupied sites) and re-
gional occupancy. Such positive abundance–occu-
pancy relationships are among the most pervasive
macroecological patterns (Gaston et al. 2000, Black-
burn et al. 2006). However, the relationship docu-
mented here was weak (r = 0.22), and strongly influ-
enced by the 'singleton effect', i.e. species recorded as
single individuals in single samples, which necessarily
have very low densities and occupancies. Across much
of the observed range of population densities, there
was substantial variation in regional occupancy
(Fig. 3). Further investigation is required to explore
more thoroughly the sources of such variation. For in-
stance, Webb et al. (2007) have shown the utility of link-
ing intra- and interspecific processes; considering the
spatial distribution of individuals within species is cer-
tainly likely to result in a better understanding of the in-
terspecific abundance–occupancy relationship (Freckle-
ton et al. 2006, Webb et al. 2007), especially if
combined with emerging statistical sampling theories
of species abundances (e.g. Green & Plotkin 2007).
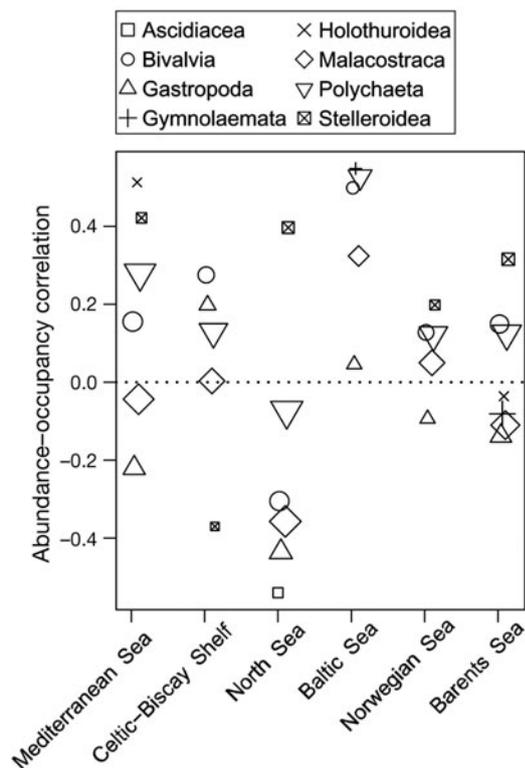What is clear, however, is that there exists substantial

Fig. 4. Summary of abundance–occupancy relationships for each class within each large marine ecosystem (LME). Each point represents the correlation between $\log_{10}$(occupancy) and $\log_{10}$(density) for a given class within a particular LME. Occupancy here is the proportion of 100 km squares occurring within an LME in which a species was recorded, and density is the mean density (ind. m^{-2}) attained by a species in samples in which it occurred within that LME. Symbols within an LME are offset slightly horizontally for clarity, and the size of a symbol is proportional to the logarithm of the number of species recorded for a given class in a given LME. The dashed horizontal line is at a correlation of 0

variation in the form and strength of abundance–occupancy relationships between geographic areas (LMEs) and higher taxa (classes); in addition, the same class can have different relationships in different areas (Fig. 4). This suggests that features of both the environment (e.g. disturbance regime) and the life histories of constituent species (e.g. larval dispersal mode) are likely to influence the form of abundance–occupancy relationships, exactly as predicted by recent population models (Freckleton et al. 2005, 2006). Given that human activity can disrupt macroecological relationships (Fisher & Frank 2004, Webb et al. 2007), it would be valuable to include measures of anthropogenic activity as covariates in future analyses; the efforts of Halpern et al. (2008) to map human impacts on marine systems would prove extremely useful here. In general, the considerable variation in abundance–occupancy relationships within the MacroBen database, combined with its wide geographic scope and the vast diversity of life

forms recorded in the database, will make it an invaluable resource for further testing of this emerging area of macroecological theory.

CONCLUSIONS

Macroecology provides a novel conceptual framework in which to analyse large-scale patterns in diversity. It has developed largely as a terrestrial subject, but the potential of marine systems to contribute to the macroecological research effort is enormous. Here, we have provided a first macroecological summary of the MacroBen database. We have shown that the European macrobenthic fauna conformed to general macroecological norms, but that considerable variation remains around certain general trends (e.g. positive interspecific abundance–occupancy relationships). The exploration of this variation, combined with the geographic and taxonomic breadth of the MacroBen database, will provide a unique opportunity to test emerging macroecological theory. Such databases will also provide one of the only means of testing hypotheses regarding the likely human impact upon macroecological patterns and relationships, and thus will provide important information on the magnitude of human effects upon marine and terrestrial ecosystems (Fisher & Frank 2004, Gaston 2004, Jennings & Blanchard 2004, Webb et al. 2007). This further illustrates the benefits to be derived from international cooperative scientific programmes such as MarBEF.

LITERATURE CITED

Blackburn TM, Gaston KJ (eds) (2003) Macroecology: concepts and consequences. Blackwell, Oxford

Blackburn TM, Gaston KJ (2006) There's more to macroecology than meets the eye. Glob Ecol Biogeogr 15:537–540

Blackburn TM, Jones KE, Cassey P, Losin N (2004) The influence of spatial resolution on macroecological patterns of range size variation: a case study using parrots (Aves: Psittaciformes) of the world. J Biogeogr 31:285–293

Blackburn TM, Cassey P, Gaston KJ (2006) Variations on a theme: sources of heterogeneity in the form of the inter-specific relationship between abundance and distribution. J Anim Ecol 75:1426–1439

Brown JH (1995) Macroecology. University of Chicago Press, Chicago, IL

Chapin FSI III, Zavaleta ES, Evlner VT, Naylor RL and others (2000) Consequences of changing biodiversity. Nature 405:234–242

Clarke A, Griffiths HJ, Linse K, Barnes DKA, Crame JA (2007) How well do we know the Antarctic marine fauna? A pre-liminary study of macroecological and biogeographical patterns in Southern Ocean gastropod and bivalve mol-luscs. Divers Distrib 12:620–632

Fisher JAD, Frank KT (2004) Abundance–distribution rela-tionships and conservation of exploited marine fishes. Mar Ecol Prog Ser 279:201–213

Foggo A, Frost MT, Attrill MJ (2003) Abundance–occupancy patterns in British estuarine macroinvertebrates. Mar Ecol Prog Ser 265:297–302

Foggo A, Bilton DT, Rundle SD (2007) Do developmental mode and dispersal shape abundance–occupancy relationships in marine macroinvertebrates? J Anim Ecol 76: 695–702

Freckleton RP, Gill JA, Noble D, Watkinson AR (2005) Large-scale population dynamics, abundance–occupancy rela-tionships and the scaling from local to regional population size. J Anim Ecol 74:353–364

Freckleton RP, Noble D, Webb TJ (2006) Distributions of habi-tat suitability and the abundance–occupancy relationship. Am Nat 167:260–275

Gaston KJ (1998) Species–range size distributions: products of speciation, extinction and transformation. Philos Trans R Soc Lond B Biol Sci 353:219–230

Gaston KJ (2003) The structure and dynamics of geographic ranges. Oxford University Press, Oxford

Gaston KJ (2004) Macroecology and people. Basic Appl Ecol 5:303–307

Gaston KJ, Blackburn TM (2000) Pattern and process in macroecology. Blackwell Science, Oxford

Gaston KJ, Blackburn TM, Greenwood JJD, Gregory RD, Quinn RM, Lawton JH (2000) Abundance–occupancy relationships. J Appl Ecol 37:39–59

Gray JS, Bjørgesæter A, Ugland K, Frank K (2006) Are there differences in structure between marine and terrestrial assemblages? J Exp Mar Biol Ecol 330:19–26

Green JL, Plotkin JB (2007) A statistical theory for sampling species abundances. Ecol Lett 10:1037–1045

Halpern BS, Walbridge S, Selkoe KA, Kappel CV and others (2008) A global map of human impact on marine ecosys-tems. Science 319:948–952

He F, Gaston KJ (2003) Occupancy, spatial variance, and the abundance of species. Am Nat 162:366–375

Hsieh CH, Reiss CS, Hunter JR, Beddington JR, May RM, Sugihara G (2006) Fishing elevates variability in the abun-dance of exploited species. Nature 443:859–862

Hunt G, Roy K, Jablonski D (2005) Species-level heritability reaffirmed. Am Nat 166:129–135

Irigoien X, Huisman J, Harris RP (2004) Global biodiversity patterns of marine phytoplankton and zooplankton. Nature 429:863–867

Jackson JBC (2001) What was natural in the coastal oceans? Proc Natl Acad Sci USA 98:5411–5418

Jennings S, Blanchard JL (2004) Fish abundance with no fish-ing: predictions based on macroecological theory. J Anim Ecol 73:632–642

Karakassis I, Eleftheriou A (1997) The continental shelf of Crete: structure of macrobenthic communities. Mar Ecol Prog Ser 160:185–196

Kerr JT, Kharouba HM, Currie DJ (2007) The macroecological contribution to global change solutions. Science 316: 1581–1584

Li WKW (2002) Macroecological patterns of phytoplankton in the northwestern North Atlantic Ocean. Nature 419: 154–157

Mackie ASY, Oliver PG, Rees EIS (1995) Benthic biodiver-sity in the southern Irish Sea. Studies in Marine Biodiver-sity and Systematics from the National Museum of Wales. BIOMÔR Reports 1. National Museum of Wales, Cardiff

Macpherson E (2003) Species range size distributions for some marine taxa in the Atlantic Ocean. Effect of latitude and depth. Biol J Linn Soc 80:437–455

May RM (1994) Biological diversity: differences between land and sea. Philos Trans R Soc Lond B Biol Sci 343:105–111

Pinheiro JC, Bates DM (2002) Mixed effects models in S and S-Plus. Springer Verlag, New York

Qian H, Ricklefs RE (2004) Geographical distribution and eco-logical conservatism of disjunct genera of vascular plants in eastern Asia and eastern North America. J Ecol 92: 253–265

Raffaelli D, Solan M, Webb TJ (2005) Do marine and terres-trial ecologists do it differently? In: Stergion KI, Browman HI (eds) Bridging the gap between aquatic and terrestrial ecology. Mar Ecol Prog Ser 304:283–289

Rao PSRS, Heckler CE (1997) The three-fold nested random effects model. J Stat Planning Inference 64:341–352

Sherman K (1991) The Large Marine Ecosystem concept: research and management strategy for living marine re-sources. Ecol Appl 1:349–360

Sokal RR, Rohlf FJ (1995) Biometry. Freeman, New York

Steele JH (1991) Can ecological theory cross the land–sea boundary? J Theor Biol 153:425–436

Stergiou KI, Browman HI (eds) (2005) Bridging the gap between aquatic and terrestrial ecology. Mar Ecol Prog Ser 304:271–307

Storch D, Sizling AL (2002) Patterns of commonness and rarity in central European birds: reliability of the core-satellite hypothesis within a large scale. Ecography 25: 405–416

Sutherland WJ, Armstrong-Brown S, Armsworth PR, Brereton T and others (2006) The identification of 100 ecological questions of high policy relevance in the UK. J Appl Ecol 43:617–627

Vanden Berghe E, Claus S, Appeltans W, Faulwetter S and others (2009) MacroBen integrated database on benthic invertebrates of European continental shelves: a tool for large-scale analysis across Europe. Mar Ecol Prog Ser 382: 225–238

Webb TJ, Gaston KJ (2003) On the heritability of geographic range sizes. Am Nat 161:553–566

Webb TJ, Gaston KJ (2005) Heritability of geographic range sizes revisited. Am Nat 166:136–143

Webb TJ, Kershaw M, Gaston KJ (2001) Rarity and phylogeny in birds. In: Lockwood JL, McKinney ML (eds) Biotic homogenization. Kluwer Academic/Plenum Publishers, New York, p 57–80

Webb TJ, Noble D, Freckleton RP (2007) Abundance–occupancy dynamics in a human dominated environment: linking interspecific and intraspecific trends in British farmland and woodland birds. J Anim Ecol 76:123–134