Facies (2005) 51: 24–32
DOI 10.1007/s10347-004-0044-3

ORIGINAL ARTICLE

Wolfgang Kiessling

Habitat effects and sampling bias on Phanerozoic reef distribution

Received: 11 January 2005 / Accepted: 9 February 2005 / Published online: 26 May 2005

Abstract Incomplete preservation, heterogeneous geographic sampling, uncertainties in palaeogeographic reconstructions and inconsistencies of reef definitions bias global reef patterns observed in the geological record. This sampling bias is added to a biological habitat area effect, which is thought to be of paramount importance for modern reefs. To evaluate the importance of sampling bias of ancient reefs, I first tested the habitat area effect and sampling bias for modern tropical reefs and then evaluated these factors for pre-Pleistocene Phanerozoic reefs. Results suggest that habitat area, although significantly affecting Phanerozoic reef patterns, is considerably less important than sampling bias. Sampling bias is more controlled by socioeconomic factors than by geological processes such as subduction of oceanic crust or sea-level fluctuations. Reefs are more likely to be sampled in rich countries, irrespective of geological and ecological controls. The numeric and geographic distribution of reefs as currently recorded in the published literature is thus probably largely artifactual, which may explain the scarcity of significant correlations between reef distribution patterns and inferred physico-chemical controls.

Introduction

Just as we shall never be able to collect all species that ever populated the earth, we are unable to discover all reefs that grew in ancient seas. The potential effects of heterogeneous preservation (meant to incorporate all sorts of bias) on our perception of ancient biodiversity patterns have received much attention in the recent literature (Peters and Foote 2001, 2002; Smith 2001; Crampton et al. 2003). These studies suggest that heterogeneities in the fossil record are so severe that even mass extinctions may be feigned (Peters and Foote 2002). No such studies are available for reefs, but there is some risk that our perception of reef distribution (meant to include numeric abundance and geographic distribution through time) is also severely affected by incomplete preservation.

The PaleoReefs project, initiated by the large literature compilation of Flügel and Flügel-Kahler (1992) and still continued by the author, aimed at a better understanding of physico-chemical and biological controls of Phanerozoic reef distribution, productivity and composition. While much has been learned from this exercise (Kiessling et al. 2002), concerns remain of how complete and unbiased our view of reef evolution really is. Nearly all previous analyses of the PaleoReefs database took the fossil record at face value without critically assessing heterogeneities of the quality of the geological record, although the existence of bias has often been noted (Flügel 1994; Kiessling 2002). Kiessling et al. (1999) and Kiessling and Flügel (2002) have demonstrated that temporal patterns of fluctuations in reef attributes are stable with increasing coverage of the published literature, but other sampling bias remains to be analysed rigorously.

This study evaluates habitat effects and sampling bias affecting the spatio-temporal distribution of Phanerozoic tropical reefs. Ancient reefs are likely to be preferentially studied where scientific and economic interests are concentrated. Recorded patterns of reef distribution might thus reflect these interests rather than ecological processes. To detail the role of this sampling bias, habitat area effects need to be considered as well. Habitat area, while representing biological constraints, may also distort our view of ancient reef distributions, because if reef distribution is largely controlled by available shelf area, other factors of interest may not be detectable (e.g., palaeoclimate, ocean chemistry). The evaluation of bias is addressed by taking modern tropical reef distribution as a reference. Thus I first discuss the habitat effects and potential sampling bias on

W. Kiessling (✉)
Institut für Paläontologie, Museum für Naturkunde,
Humboldt-Universität Berlin,
Invalidenstr. 43,
10115 Berlin, Germany
e-mail: wolfgang.kiessling@museum.hu-berlin.de
Tel.: +49-30-2093-8576
Fax: +49-30-2093-8868

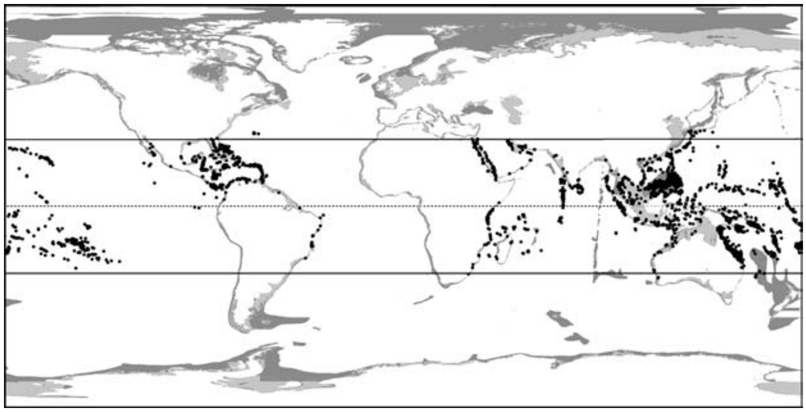

Fig. 1 Global distribution of modern zooxanthellate coral reefs (*black dots*), area with less than 100 m water depth (*light grey*) and continental crust (*dark grey*). *Solid horizontal lines* are at 30° latitude, *dashed line* marks the equator. Sources stated in text

modern tropical coral reef distribution and go on to compare these results with ancient reef distributions.

Data and methods

The PaleoReefs database (Kiessling and Flügel 2002) was used to analyse palaeogeographic reef patterns. Data on 3,340 pre-Pleistocene Phanerozoic reef complexes are currently available in PaleoReefs. Shelf areas of ancient time intervals were measured with ArcView Spatial Analyst using the palaeogeographic reconstructions of Jan Golonka as published in the book Phanerozoic Reef Patterns (Kiessling et al. 2002). Both the PaleoReefs database and the palaeogeographic reconstructions are available online (http://193.175.236.205/paleo; id = paleo, password= reefs). Although a finer stratigraphic subdivision of the reef record is possible, the analyses in this paper are based on supersequences with an average duration of 17 million years (Golonka and Kiessling 2002).

Reefbase, a compilation of modern tropical coral reefs (http://www.reefbase.org/), forms the primary reference for my analysis of modern reef distributions (download 1 December 2004). Reefbase currently lists more than 10,000 tropical and subtropical reefs, and besides reefs also includes data on non-reef coral communities (263 entries). Also available at Reefbase is information on shelf area and reef area per country. Although not complete (70 countries with shelf area, 107 countries and dependent regions with reef area), this provides an independent source of information for tests between potential habitat area and reef area.

Information on bathymetric data of modern oceans was extracted from the geophysical data system GEODAS (http://www.ngdc.noaa.gov/mgg/geodas/geodas.html) using data from the ETOPO-2 project. The bathymetric data for the tropics and subtropics stem from just one source (Smith and Sandwell 1997). I utilised a custom grid of 10 min geographic resolution and a full meters bathymetric precision, which was imported into and analysed with ArcView and ArcView Spatial Analyst. The potential habitat area for reefs was defined between 0 and 99 m water depth, referred to as shallow-water area in this paper. Another source of information for habitat area was used to permit a more straightforward comparison with ancient settings, where bathymetry is often poorly resolved. We usually have information on continental areas, where marine sediments have been deposited. These areas are marked as continental shelf in palaeogeographic reconstructions. To achieve a similar measure for the Recent, I have used Scotese's shapefile of continental plates (pers. comm., 2001) and subtracted from this the area of exposed land with ArcView's "create doughnut polygon" feature (Fig. 1).

Both PaleoReefs and Reefbase collect data on various reefal structures, but Reefbase only comprises tropical to subtropical coral reefs, whereas non-tropical reef types are included in PaleoReefs. In PaleoReefs individual data entries have a 20 km minimum distance for each time line; thus reef numbers as reported in PaleoReefs represent reef sites rather than individual reefs. There is no such restriction for Reefbase. To better compare patterns in both databases, the following adjustments were performed. In PaleoReefs non-tropical reef types (reefs that are distinctly different from reefs in the tropics and occur in high palaeolatitudes) were excluded, which leaves 3,267 pre-Pleistocene Phanerozoic reefs. The data from Reefbase have been filtered to exclude non-reef coral communities and to lump data from closely spaced reefs to a minimum spacing of 20 km. This leaves 2,332 modern reefs, which are then also better referred to as reef sites. This reduced Reefbase (Reefbase-R) is the source of all analyses of modern reef distribution, if not otherwise indicated.

Socioeconomic data for countries were extracted from additional data sources. For economical data I have relied on the Energy Information Administration report at http://www.eia.doe.gov/emeu/iea/popgdp.html and extracted the gross domestic product (GDP) of 2002. Other data such as land area, population, literacy of population and geographic coordinates were taken from the CIA factbook at http://www.cia.gov/cia/publications/factbook/.

Statistical tests in this paper are limited to simple correlation tests. Pearson correlations were applied when data are normally distributed; otherwise Spearman-Rho correlations were used. Data with large ranges (e.g., shelf areas per country) have been log-transformed prior to analysis to achieve a normal distribution of the data. Proportional data were prepared for analysis using the standard logit transformation: $\log(x/[1-x])$.

Modern reef distribution

The distribution of modern zooxanthellate coral reefs is well constrained. Compared with our ignorance concerning the number of species inhabiting our planet, we have a fairly good knowledge of the geographic distribution of reefs and the total oceanic area covered by coral reefs. Recent estimates vary between 284,300 km^2 (Spalding et al. 2001) and 345,000 km^2 (Vecsei 2004). The differences in these estimates is less due to uncertainties of information but rather due to differences in the definition of coral reef area. The absolute number of reefs is less clear, but this is probably also due to inconsistencies in separating individual reefs rather than reflecting gaps in our knowledge. The number of reefs recorded in Reefbase has risen from about 7,000 to more than 10,000 between 2001 and 2004. This 43% increase in total reef numbers compares to an increase of just 15% in Reefbase-R over the same period, implying that comparatively few new reef sites have been added.

The spatial distribution of modern reefs is determined by physico-chemical factors and available habitat area (Fig. 1). Ecological constraints of reef corals limit the latitudinal distribution of coral reefs to the tropics and lower subtropics and to oligotrophic to mesotrophic regions. The currently known latitudinal boundaries are at 32° S and 34° N. Within this latitudinal range, reefs are non-randomly distributed. A plot of reef numbers per degree of latitude shows that relatively few reefs are situated around the equator and the maximum concentration of reefs is between 16° and 22° in the Southern Hemisphere and between 5° and 15° in the Northern Hemisphere (Fig. 2). Reef numbers drop off markedly at 24°S and 29°N, before the actual limit of shallow-water reef growth.

To test if habitat area or other ecological constraints are responsible for these patterns, I have analysed cross-correlations between reef numbers and areas of total and shallow shelf in latitudinal bands ranging from 34° N to 34° S. The correlation between the number of reefs and those measures of habitat area in 1-degree latitudinal intervals is not perfect, but statistically significant (Fig. 3; Table 1). Correlations are generally greater for Reefbase-R

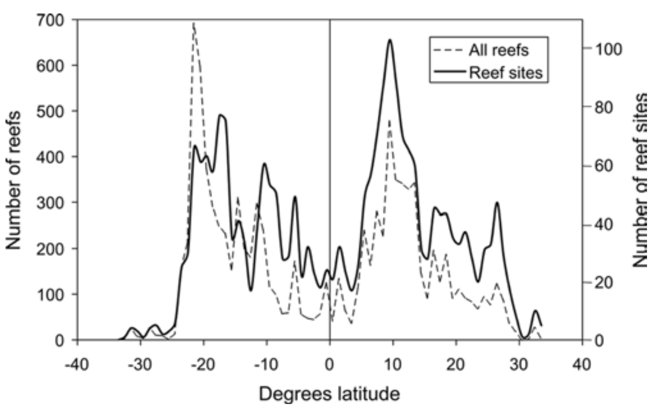

Fig. 2 Latitudinal distribution of modern coral reefs. Counts of reefs per 1 degree of latitude based on the full dataset in Reefbase (all reefs, *dashed line*) and the reduced dataset (reef sites, *solid line*)

Table 1 Correlation between reef abundance and habitat area in one degrees latitudinal bands

Reef abundance	Shelf area (<100 m depth)	Shelf area (coarse)
Pearson correlation of raw data		
Reefbase[a]	0.15	0.38[d]
Reefbase-R[a]	0.41[d]	0.59[d]
PaleoReefs[b]	N.A	0.46[d]
Pearson correlation of detrended data		
Reefbase[a]	0.12	0.28[c]
Reefbase-R[a]	0.23	0.32[d]
PaleoReefs[b]	N.A	0.14

[a]Data within 34° North and South
[b]Data within 40° North and South
[c]Correlation is significant at the 0.05 level (2-tailed)
[d]Correlation is significant at the 0.001 level (2-tailed)
N.A. = not applicable

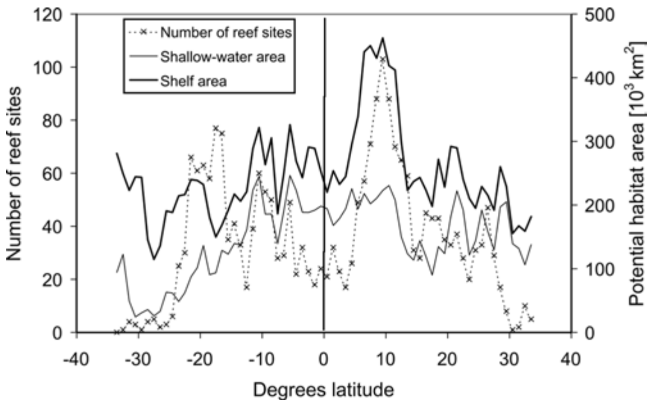

Fig. 3 Latitudinal distribution of modern coral reef sites compared to distribution of potential habitat area. Shallow-water area refers to submerged area with less than 100 m depth. Shelf area refers to area of submerged continental plates
Table 2

underlining the need for a homogeneous protocol to define individual reefs. While correlations of raw data are relatively high, these values are not reliable, because there are strong spatial autocorrelations in the data, that is, values in one latitudinal interval predict values in the next interval to a certain degree. Similar to the approach in time series these autocorrelations can be eliminated by generalised differencing, where the first-order autocorrelation coefficient is subtracted from the dataset (McKinney and Owen 1989). When autocorrelations are accounted for, significant correlations are only evident between the generalised differences of reef numbers and total shelf area (Table 1). Changes in shelf area thus predict changes in reef numbers, but based on R^2, only 10% of the variance in changes of reef numbers are explained by changes in shelf area. The correlations of detrended values are somewhat better when the analysis is restricted to the tropics (less than 24° latitude) suggesting that other ecological constraints become more important in higher latitudes. The same is obviously true for the equatorial region. When the analysis is restricted to areas between 5° and 24° latitude, the correlation coefficients become larger and the correlation

Table 2 Relationships between modern reef and shelf areas by country

	Shelf areas within 34 degrees latitude	Shelf areas within 24 degrees latitude
Reef areas	*r*=0.57	*r*=0.53
	N=54	*N*=42
Reef numbers	*r*=0.62	*r*=0.59
	N=50	*N*=38

Pearson correlation of log-transformed values
All *p*-values <0.001
Source of data: http://www.reefbase.org

between reef abundance and habitat area is significant for both total shelf and shallow shelf. Thus the number of reef sites is controlled by available potential habitat area but other ecological factors such as temperature and nutrient concentrations may be more important. Temperature constraints may blur the correlation in higher latitudes, while both nutrient concentrations and temperature may be responsible for the unusually low abundance of reefs near the equator. Elevated nutrient concentrations near the equator have also been held responsible for a gap in the distribution of Recent carbonate banks (Vecsei 2003).

To avoid the problem of spatial autocorrelations, I have tested the correlation between reef abundance and shelf area by country. The first test included only countries with shelf areas completely within the global latitudinal limit of reef growth (less than 34° latitude). A second test was performed to only include countries whose shelves are completely within the tropics (less than 24° latitude). Due to log-transformations, this test implicitly excludes shelves without reefs. Highly significant correlations are evident between shelf area and two measures of reef abundance: reef numbers and reef area (Table 2). Up to 33% of the variance in reef area and up to 38% of the variance in the number of reef sites is explained by shelf area.

In summary, two independent tests suggest that the size of available habitat area has a significant effect on geographic reef distribution. The importance of this habitat effect is not completely clear, because results differ somewhat between tests and metrics. The surprisingly small correlation between shallow-water shelf area and reef abundance in latitudinal bands is partially attributable to problems with spatial autocorrelations, but the highest distortion is probably due to the large proportion of reefs (24%) occurring outside shallow-water area as detected by grid resolution in ArcView (see chapter Data and methods).

Even for the highest correlations between reef abundance and habitat area (Table 2), the habitat effect is less than expected from a previous analysis, which suggested that habitat area is the major determinant of reef diversity (Bellwood and Hughes 2001). We may conclude that additional ecological factors are required to explain modern reef distribution. These factors can be extracted when the habitat-area effect is accounted for.

Sampling bias is not expected to be significant for modern shallow-water reefs. Although the most detailed biological data stem from shelves situated around rich countries such as Australia, Japan and the United States, reefs are mapped and monitored at global scales by the Global Coral Reef Monitoring Network (GCRMN). A simple test was performed to check for sampling bias. Reef numbers and reef area as provided in Reefbase were normalised for shelf area and tested for correlations with GDP, which was also normalised for shelf area. None of the parameters has any significant correlations (based on Spearman-Rho, r_S this metric used, because data are not normally distributed), which supports the view that sampling bias is negligible for modern tropical coral reefs.

Ancient reef distribution

Global patterns of numeric abundance, reefal carbonate production and palaeogeographic distribution of Phanerozoic reefs have been discussed in several reviews in the last few years (Kiessling et al. 1999; Kiessling et al. 2000; Kiessling 2001; Kiessling et al. 2002). The results suggest large temporal variance in all these traits, with carbonate production fluctuating most, followed by reef numbers and latitudinal range.

Increasing sampling over the years has resulted in little change of the overall pattern (Fig. 4), which means that the probability of sampling reefs in a particular time interval is fairly well represented by the current database. However, the total number and palaeogeographic distribution of ancient reefs must be severely incomplete, if we assume that reefs were at least at some times in the past as proliferating as they are today. The total number of pre-Pleistocene Phanerozoic reefs is only 40% larger than the number of modern reefs, although the Phanerozoic is approximately 300 times longer than the Pleistocene, which is essentially the foundation period of modern reefs.

The two major questions relevant for this study are now: (1) What is the magnitude of and what are the reasons for sampling bias and (2) how much of the variance is due

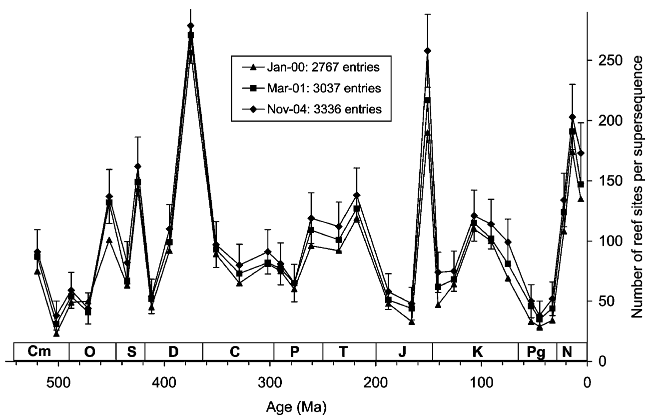

Fig. 4 Number of reef sites per supersequence recorded in the PaleoReefs database. The three curves represent numbers at different stages of database development. Error bars in the Nov-04 curve are binomial errors (95% confidence intervals) calculated from percentage values over the entire Phanerozoic

to variance in available habitat area? These two questions cannot be treated in isolation because habitat and sampling effects are intimately linked in ancient reefs.

Sampling bias

There are two main motivations for the exploration of ancient reefs. One is scientific interest, largely driven by theoretical questions of ecological and evolutionary processes. The second is economic interest, largely driven by the need of energy and raw materials. Both issues tend to bias our perception of ancient reef distributions. Reef data will preferentially become available from places and time intervals (collectively called "areas" subsequently) where scientific interest is concentrated and the economic value of reefs is high (e.g., they are rich in hydrocarbons). Ideally, theoretical biogeologists will focus their research on areas where the scientific benefit is greatest. Bias would then be minimal, because research would focus on previously unexplored areas. However, there is danger that biogeologists will rather go where access to new information is eased by available infrastructure. If this is the case, sampling bias is expected to be large. Spatio-temporal bias by economic interest is expected to be strongest when there is a strong spatio-temporal heterogeneity in economic value.

When plotted on a present-day geographic map, the distribution of ancient reefs indeed suggests a severe sampling bias, with considerable concentrations of reefs in countries with a long tradition in geosciences (e.g., Europe and North America) and regions where the economic potential of reefs is great (e.g., West Canada, Polar Urals, South China; Fig. 5).

Similar to modern reefs, the magnitude of sampling bias is best explored on a per country basis. I have summed reef numbers per country in PaleoReefs and checked for correlations with geographic and socioeconomic factors. To reduce noise, only countries with at least 10 reefs (N=60) were included in the analyses. The number of reef sites per country is highly skewed, with reef occurrences in 125 countries (plus occurrences in Antarctica and other regions outside official country borders), but 50% of all reefs are from just nine countries (United States, Russia, Canada, China, Italy, France, Spain, Indonesia, Germany). This is not surprising because there is a similar skew in the distribution of land area among countries. Indeed, there is a highly significant correlation between land area and the number of reef sites recorded in each country (r_S=0.395, P=0.002). This correlation, however, is less than expected. Although Recent land area does not necessarily correlate with ancient shelf area, this may indicate that factors other than sampling area may be more relevant. The observation that the countries hosting the most reefs are either wealthy and/or have an unusual high percentage of reefs with hydrocarbon reservoir quality, suggest that socioeconomic factors might represent a significant bias in the record of ancient reefs.

Among the tabulated socioeconomic factors such as gross domestic product (GDP), population size, percent literacy and per capita GDP, the first two factors show significant Spearman-Rho correlations with reef numbers and both explain more of the variance in reef numbers than land area. To omit the effect of land area on the correlations, I have normalised all variables for land area (Fig. 6). The resulting correlations are very high for both population density *versus* reef density (r_S=0.614, $P<0.001$) and GDP density *versus* reef density (r_S=0.735, $P<0.001$). These results suggest that socioeconomic factors such as population density and especially economic productivity introduce a severe bias in reef distributions. Because reefs are more likely to be recorded in wealthy countries with high population densities, the palaeolatitudinal distribution of reefs will be biased by the plate tectonic movement of these countries and reefs will be concentrated in time intervals when the marine geological record is good in these countries.

The economic value of reefs does not seem to introduce a significant sampling bias. Neither the logit-transformed proportions of reefs with reservoir potential nor the proportions of subsurface reefs have any correlation with the number of reef sites or preserved reef volumes on a per country basis. Thus although economic interest may produce regional concentrations in the record of reefs, this effect is too small to produce a severe bias in large-scale reef patterns.

Another sampling bias, which cannot be avoided by focusing future research on previously underexplored areas, is given by the large difference in crustal position of modern and ancient reefs. 37±2% of all modern coral reefs are

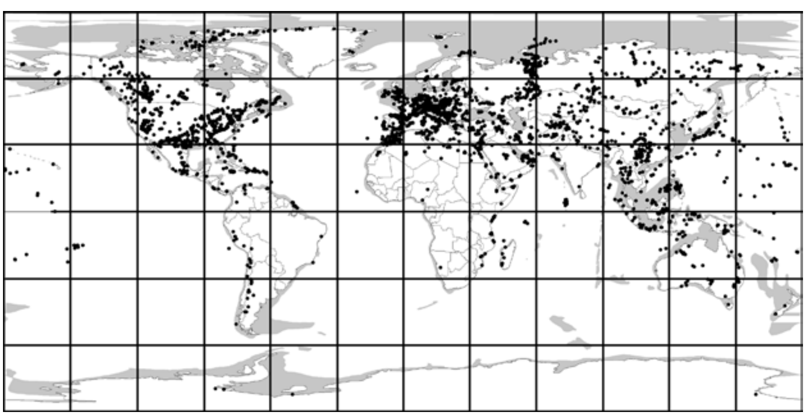

Fig. 5 Distribution of recorded pre-Pleistocene Phanerozoic reef sites plotted on a present day geographic map. Note strong concentration of reef sites in Europe

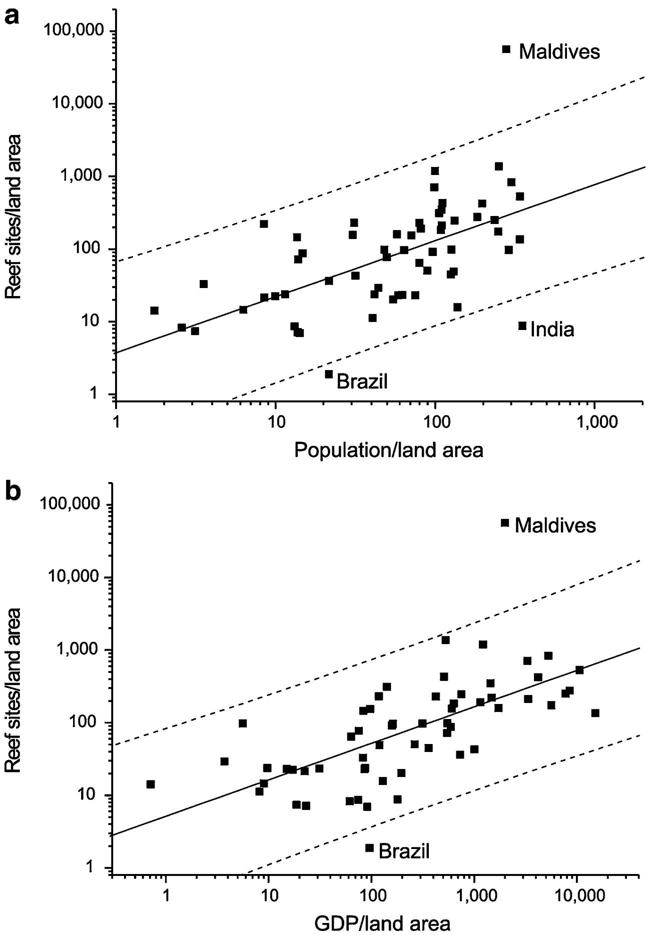

Fig. 6 Correlation between the density of ancient reef sites and socioeconomic factors per unit area. **a** Population per km^2 land area *versus* reef sites per 10^6 km^2 land area. **b** Gross domestic product in Billion US-$ per 10^6 km^2 land area versus reef sites per 10^6 km^2 land area. *Solid line* indicates least-squares regression; *dashed lines* indicate 95% prediction bands. Significant outliers in both plots are specified. Note that the pre-Pleistocene reef record of the Maldives and Brazil is exclusively (Maldives) or nearly so (Brazil) from subsurface exploration

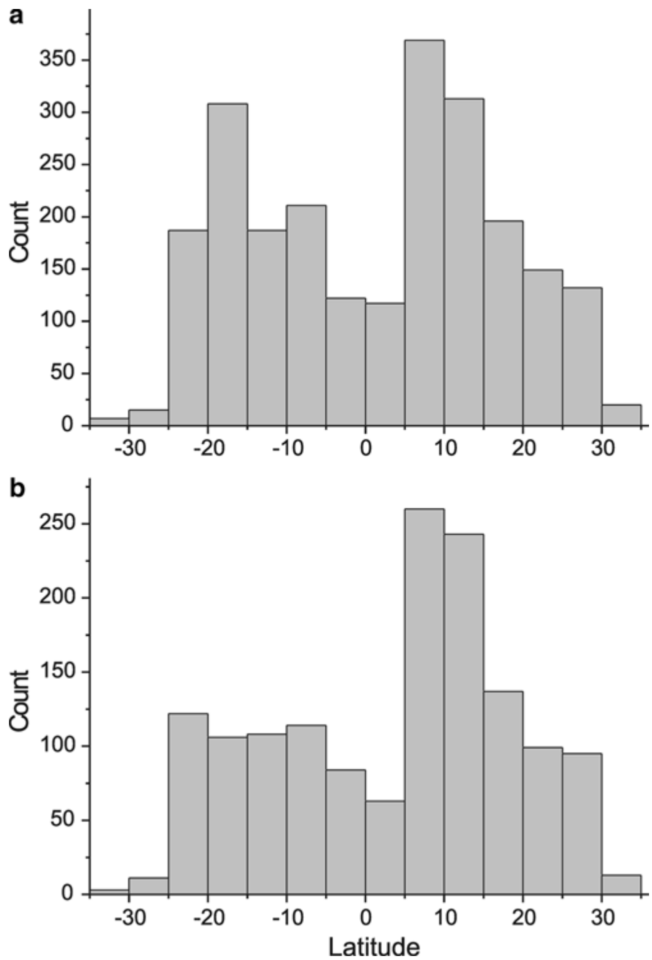

Fig. 7 Histograms of modern reef distribution in five degrees latitudinal intervals. **a** Counts for all reefs in Reefbase-R. **b** Count for reefs excluding reefs on oceanic crust. Note that although there is some change, especially in the Southern Hemisphere, the overall latitudinal reef distribution is well depicted by continental reefs

situated in oceanic regions, while the proportion of reefs outside continental shelf area in PaleoReefs is only 5±1% with all ancient reefs on oceanic crust being from the Cretaceous and Cenozoic. Although the low value in PaleoReefs is biased because many older oceanic reefs were later accreted to continental crust, the great majority of ancient oceanic reefs are likely to have been lost from the geological record by subduction. If the proportion of oceanic reefs was similar in the past as it is today, we would constantly miss one third of the reefs in any pre-Cretaceous time interval. For reef numbers and carbonate production, this would be no severe problem because the proportion of missing reefs could simply be added to these older time intervals. For geographic distribution, however, the bias is difficult to assess. The latitudinal distribution of modern reefs is little affected when all oceanic reefs are dropped from the analysis (Fig. 7). There is a slight shift towards the Northern Hemisphere with the mean latitude of reefs increasing from 1.5 to 3.7 degrees. But the latitudinal boundaries of reef distribution are as stable as the equatorial drop in reefs numbers. This result is encouraging because it raises the faith that physico-chemical factors presumably driving the geographic distribution of reefs can be extracted from just looking at the continental record of reefs.

Habitat effects

The potentially available habitat for ancient reefs cannot be quantified with the same precision as for modern reefs. Only the area of continental shelf can be reliably estimated based on global palaeogeographic reconstructions (Golonka 2002). As discussed under modern reef distributions, this metric, although only a rough approximation of ancient potential habitat area, is at least as useful as more sophisticated proxies of shallow-water area. The advantage of these reconstructions is that they are palinspastic, interpretative maps, which are not affected by sampling bias. Equivalent to modern reefs, two tests were performed on the dependency of reef distribution and shelf area.

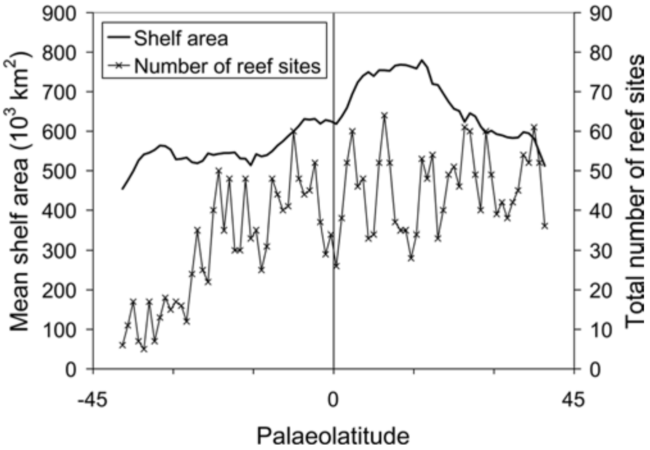

Fig. 8 Palaeolatitudinal distribution of tropical pre-Pleistocene Phanerozoic reefs and shelf area. Curves represent cumulative values summed for 32 Phanerozoic supersequences

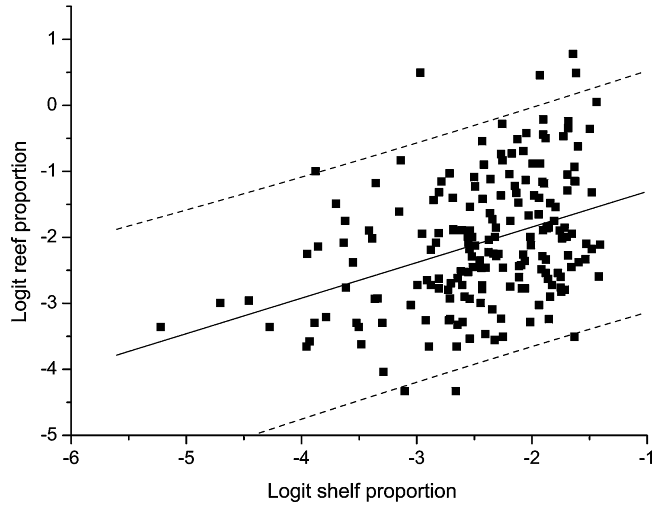

Fig. 9 Correlation between the proportion of Phanerozoic reef sites and the proportion of shelf area in 30 degrees grid cells. Cells combined from 32 Phanerozoic supersequences for palaeolatitudes of up to 30 degrees. *Solid line* indicates least-squares regression; *dashed lines* indicate 95% prediction bands

The first test assesses correlations between shelf areas and reef numbers in latitudinal intervals. As for modern reefs, I have tallied shelf areas and reef numbers into palaeolatitudinal bands. The measurements represent cumulative values of all 32 supersequences to maximise sample size. Included in the tests were only tropical reef types located between 40° North and South where ancient tropical reefs are concentrated (Fig. 8). Correlation tests were performed at different sample resolutions with 1-degree, 2-degrees and 5-degrees latitudinal intervals. All tests show significant correlations of the raw data but no significant correlations of detrended data. Hence although the number of reef sites tends to be higher in latitudinal bands with greater shelf area, changes in shelf area do not correspond to changes in reef abundance. The test results are the same if proportional rather than absolute measures are compared and if data from all time slices are treated in isolation.

A second test is similar to the by-country test for modern reefs. Adjustments were necessary because shelf areas per country do not make sense owing to the common separation into multiple tectonic plates. I have measured the shelf area in 30° palaeogeographic grid cells and tested this shelf size for correlations with reef numbers and carbonate production in the same grids. The test includes measurements for all 32 supersequences in PaleoReefs (Golonka and Kiessling 2002) but excludes areas from latitudes greater than 30° to minimise the influence of climatic factors. Grid cells with less than 3 reefs were excluded from the analysis to minimise noise. A total of 189 cells passed the filter. All values were transformed to proportional data within time slices prior to analysis, in order to balance for different absolute reef numbers and shelf sizes. The correlations between logit-transformed values are significant between shelf area and both the number of reef sites (r_S=0.38, P<0.001) and reef volume (r_S=0.16, P=0.031) within grid cells. Because there are no significant autocorrelations in the data, the correlation test is straightforward and indicates a significant dependence especially of recorded reef numbers on primary shelf area (Fig. 9). I have then also included shelf areas up to 60° latitude with a total of 254 grid cells. Surprisingly, correlations become even stronger with r_S=0.43, P<0.001 for reef numbers and shelf area, and r_S=0.18, P=0.004 for reef volume and shelf area.

In summary, there are significant correlations between potential habitat area and recorded reef abundance. The correlations, however, are much weaker than the correlation with socioeconomic factors leaving habitat area of only secondary importance. As there is no reason to assume that habitat area effects were less important in the past than they are today, I suspect that sampling bias due to socioeconomic factors may blur the correlations.

Discussion

The analyses performed in this paper suggest that habitat area is a significant factor controlling modern and ancient reef patterns. In contrast to modern reefs, however, the record of ancient reefs is strongly affected by sampling bias, which exceeds the influence of habitat area. Socioeconomic factors seem to introduce the largest bias. Among all factors tested in this paper, the economic productivity of countries as measured by gross domestic product (GDP) is the single most important control on Phanerozoic reef patterns, explaining 54% of the variance in reef numbers.

There are of course many more potential biases, which were not tested in this paper. Intuitively important are changes in continental freeboard (affecting both habitat and sampling area) and the area of exposed sedimentary rocks (affecting sampling area). Previous analyses have shown that eustatic sea level, continental freeboard are indeed significantly correlated with reef numbers and carbonate production (Kiessling 2002) but correlation coefficients are far smaller (maximum r_S=0.41, P=0.02 between preserved reef volume and continental freeboard) than for socioeconomic factors.

The area or volume of exposed sedimentary rocks is long known as a significant bias in preserved biodiversity (Raup 1976) and continues to be the main factor against which the quality of the fossil record is tested (Smith 2001; Crampton et al. 2003). There is indeed all reason to assume that outcrop area is one of the major determinants of preservation of both biodiversity and reef structures. However, previous analyses are either limited in spatio-temporal scope or have a stratigraphic resolution far coarser than applicable for reefs. The currently most detailed global dataset stems from the research group of Ronov (Ronov et al. 1984, 1989). I have extracted the volume and area of marine carbonates from this dataset into epochs (see also Walker et al. 2002) and found no significant correlations with reef abundance data neither with raw data nor with detrended data. However, the outcome is equivocal because these data refer to calculated areas and volumes and do not necessarily reflect exposure areas. The global dataset with actual sedimentary outcrop areas used by Raup (1976) also shows limited correlations with reef abundance data. The raw data suggest a significant correlation between outcrop area per million years and the number of reef sites per million years (r_S=0.68, P=0.032) but detrended values show no correlation. Although the outcome of this test is also equivocal, owing to the limited sample size (data parsed into geological periods), the results of both tests imply that the preserved area of sedimentary rocks has a more limited effect on recorded reef abundance than on preserved biodiversity. These results also suggest that the sampling bias introduced by outcrop area and volume is far smaller than the bias of socioeconomic factors.

Conclusions

Ancient reef patterns as recorded in the published literature are affected by several factors that tend to mask the effect of physico-chemical and ecological factors, which are the focus of most studies. The single most important factor affecting the distribution of ancient reefs, much more important than any other currently testable bias, is economic wealth as expressed by the gross domestic product of countries. The great majority of ancient reef data stem from wealthy countries and there is no reason to assume that these countries actually host most of the reefs. The concentration of reef research in wealthy countries leads to a spatio-temporal reef distribution that is perhaps artificial and it is thus no surprise that significant correlations are scarce between earth system parameters and absolute values of reef abundance, reefal carbonate production and latitudinal range (Webb 1996; Kiessling 2002). Although primarily affecting (palaeo)geographic distributions and numeric abundance patterns of reefs, the artificial concentrations of reef data may also bias our perception of reef evolution as a hole. Because pre-Pleistocene Phanerozoic reefs were usually geographically much less uniform in composition and architecture than modern reefs, there is a risk that even measures that are based on proportional rather than absolute data (e.g., guild structure and petrography) might be distorted.

This conclusion is related to a previous study, which states that recorded biodiversity is largely determined by palaeontological interest rather than by true species richness (Sheehan 1977). However, the observation that this interest is largely driven by economic factors is frustrating. Although geologists become increasingly active in remote areas and so-called underdeveloped countries, we are far from achieving a homogeneous geographic distribution of biogeological effort. To overcome this geographic sampling bias, one could apply statistical resampling techniques similar to the ones developed for palaeo-biodiversity studies (Alroy et al. 2001; Bush et al. 2004). However, severe heterogeneities in geographic coverage continue to cause problems even with the most sophisticated statistical techniques (Alroy et al. 2001). It would thus be preferable to achieve a more homogenous geographic sampling of reefs, by focusing research on poorly explored areas. This advice is well in line with Erik Flügel's policy to encourage scientists to publish their results from underexplored areas (e.g., central Asia, Russia) in FACIES. A significant intensification of this policy, also in other journals and funding agencies, will be necessary to even out the current heterogeneities in sampling intensity.

Acknowledgements This study is devoted to my teacher Erik Flügel, whose enthusiastic support made me change my primary research focus from micropalaeontology to reef studies and palaeontological databases. The broad and courageous view of Erik Flügel got both of us into the study of large-scale trends in reef ecosystems. Erik's enthusiasm was always mixed with concern about potential biases in the record of ancient reefs. I hope to have at least partially responded to his concerns with this paper. I thank André Freiwald and an anonymous reviewer for corrections.

References

Alroy J, Marshall CR, Bambach RK, Bezusko K, Foote M, Fürsich FT, Hansen TA, Holland SM, Ivany LC, Jablonski D, Jacobs DK, Jones DC, Kosnik MA, Lidgard S, Low S, Miller AI, Novack-Gottshall PM, Olszewski TD, Patzkowsky ME, Raup DM, Roy K, Sepkoski JJ Jr., Sommers MG, Wagner PJ, Webber A (2001) Effects of sampling standardization on estimates of Phanerozoic marine diversification. Proc Natl Acad Sci (USA) 98:6261–6266

Bellwood DR, Hughes TP (2001) Regional-scale assembly rules and biodiversity of coral reefs. Science 292:1532–1535

Bush AM, Markey MJ, Marshall CR (2004) Removing bias from diversity curves: the effects of spatially organized biodiversity on sampling-standardization. Paleobiology 30:666–686

Crampton JS, Beu AG, Cooper RA, Jones CM, Marshall B, Maxwell PA (2003) Estimating the rock volume bias in paleobiodiversity studies. Science 301:358–360

Flügel E (1994) Pangean shelf carbonates: Controls and paleoclimatic significance of Permian and Triassic reefs. In: Klein GD (ed) Pangea: paleoclimate tectonics and sedimentation during accretion zenith and breakup of a supercontinent. Geol Soc Amer Spec Pap 288:247–266

Flügel E, Flügel-Kahler E (1992) Phanerozoic reef evolution: basic questions and data base. Facies 26:167–278

Golonka J (2002) Plate-tectonic maps of the Phanerozoic. In: Kiessling W, Flügel E, Golonka J (eds) Phanerozoic reef patterns. SEPM Spec Publ 72:21–75

Golonka J, Kiessling W (2002) Phanerozoic time scale and definition of time slices. In: Kiessling W, Flügel E, Golonka J (eds) Phanerozoic reef patterns. SEPM Spec Publ 72:11–20

Kiessling W (2001) Paleoclimatic significance of Phanerozoic reefs. Geology 29:751–754
Kiessling W (2002) Secular variations in the Phanerozoic reef ecosystem. In: Kiessling W, Flügel E, Golonka J (eds) Phanerozoic reef patterns. SEPM Spec Publ 72:625–690
Kiessling W, Flügel E (2002) PaleoReefs - a database on Phanerozoic reefs. In: Kiessling W, Flügel E, Golonka J (eds) Phanerozoic reef patterns. SEPM Spec Publ 72:77–92
Kiessling W, Flügel E, Golonka J (1999) Paleoreef maps: Evaluation of a comprehensive database on Phanerozoic reefs. AAPG Bull 83:1552–1587
Kiessling W, Flügel E, Golonka J (2000) Fluctuations in the carbonate production of Phanerozoic reefs. In: Insalaco E, Skelton PW, Palmer TJ (eds) Carbonate platform systems: components and interactions. Geol Soc London Spec Publ 178:191–215
Kiessling W, Flügel E, Golonka J (2002) Phanerozoic Reef Patterns. SEPM Spec Publ 72:775 pp
McKinney ML, Owen CW (1989) Causation and nonrandomness in biological and geological time series: temperature as a proximal control of extinction and diversity. Palaios 4:3–15
Peters SE, Foote M (2001) Biodiversity in the Phanerozoic: a reinterpretation. Paleobiology 27:583–601
Peters SE, Foote M (2002) Determinants of extinction in the fossil record. Nature 416:420–424
Raup DM (1976) Species diversity in the Phanerozoic: an interpretation. Paleobiology 2:289–297
Ronov A, Khain V, Seslavinski A (1984) Atlas of lithological paleogeographical maps of the world: Late Precambrian and Paleozoic of the continents. USSR Acad Sci Leningrad, 70 pp
Ronov A, Khain V, Balukhovski A (1989) Atlas of lithological paleogeographical maps of the world: Mesozoic and Cenozoic of the continents. USSR Acad Sci Leningrad, 79 pp
Sheehan PM (1977) Species diversity in the Phanerozoic: a reflection of labor by systematists? Paleobiology 3:325–329
Smith AB (2001) Large-scale heterogeneity of the fossil record: implications for Phanerozoic biodiversity studies. Phil Trans R Soc London, B 356:351–367
Smith WHF, Sandwell DT (1997) Global sea floor topography from satellite altimetry and ship depth soundings. Science 277:1956–1962
Spalding MD, Ravilious C, Green EP (2001) World atlas of coral reefs. Univ California Press, Berkeley, 424 pp
Vecsei A (2003) Nutrient control of the global occurrence of isolated carbonate banks. Int J Earth Sci 92:476–481
Vecsei A (2004) A new estimate of global reefal carbonate production including the fore-reefs. Global Planet Change 43:1–18
Walker LJ, Wilkinson BH, Ivany LC (2002) Continental drift and Phanerozoic carbonate accumulation in shallow-shelf and deep-marine settings. J Geol 110:75–87
Webb GE (1996) Was Phanerozoic reef history controlled by the distribution of non-enzymatically secreted reef carbonates (microbial carbonate and biologically induced cement)? Sedimentology 43:947–971