Palaeontology The Palaeontological Association www.palass.org
[Palaeontology, Vol. 55, Part 6, 2012, pp. 1155–1175]

COMPLETENESS OF THE FOSSIL RECORD AND THE VALIDITY OF SAMPLING PROXIES AT OUTCROP LEVEL

by ALEXANDER M. DUNHILL[1]*, MICHAEL J. BENTON[1], RICHARD J. TWITCHETT[2] *and* ANDREW J. NEWELL[3]

[1]School of Earth Sciences, University of Bristol, Wills Memorial Building, Queen's Road, Bristol BS8 1RJ, UK; e-mails: alex.dunhill@bristol.ac.uk; mike.benton@bristol.ac.uk
[2]School of Geography, Earth and Environmental Sciences, University of Plymouth, Drake Circus, Plymouth PL4 8AA, UK; e-mail: richard.twitchett@plymouth.ac.uk
[3]British Geological Survey, Maclean Building, Crowmarsh Gifford, Wallingford OX10 8BB, UK; e-mail: ajn@bgs.ac.uk
*Corresponding author.

Typescript received 13 May 2011; accepted in revised form 16 December 2011

Abstract: Most studies of the adequacy of the fossil record have been carried out at a global or continental scale, and they have used sampling proxies that generally do not incorporate all aspects of sampling (i.e. rock volume, accessibility, effort). Nonetheless, such studies have identified positive correlations between apparent diversity and various sampling proxies. The covariation of fossil and rock record signals has been interpreted as evidence for bias or for a common cause, such as sea level change, or as evidence that the signals are in some ways redundant with each other. Here, we compare a number of proxies representing the three main aspects of sampling, (1) sedimentary rock volume, (2) rock accessibility and (3) worker effort, with palaeodiversity in a geographically and stratigraphically constrained data set, the marine Lower Jurassic outcrop of the Dorset and East Devon Coast. We find that the proxies for rock volume and accessibility do not correlate well with the other sampling proxies, nor with apparent diversity, suggesting that the total amount of sedimentary rock preserved does not influence apparent diversity at a local scale, that is, the rock record at outcrop has been sampled efficiently. However, we do find some correlations between apparent diversity and proxies for worker effort. The fact that the proxies do not correlate significantly with each other suggests that none can be regarded as an all-encompassing sampling proxy that covers all aspects of bias. Further, the presence of some correlations between sampling proxies and diversity most probably indicates the bonanza effect, as palaeontologists have preferentially sampled the richest rock units.

Key words: diversity, sampling, proxy, GIS, Dorset, Lower Jurassic.

Many recent publications have highlighted a significant positive correlation between palaeodiversity and various proxies for sampling (Raup 1972; Sheehan 1977; Peters and Foote 2001; Smith 2001; Smith *et al.* 2001; Peters and Foote 2002; Crampton *et al.* 2003; Smith and McGowan 2005, 2007; Wang and Dodson 2006; Fröbisch 2008; Barrett *et al.* 2009; Butler *et al.* 2009, 2011; Wall *et al.* 2009; Benson *et al.* 2010; Mannion and Upchurch 2010). The covariation between palaeodiversity and sampling proxies has led to four possible interpretations that the fossil record is significantly influenced by (1) bias driven by the amount of sedimentary rock available for sampling (Raup 1972, 1976; Peters and Foote 2001, Smith 2001, 2007; Peters and Foote 2002; Smith and McGowan 2005, 2007; Fröbisch 2008; Barrett *et al.* 2009, Wall *et al.* 2009), (2) variable human collecting effort expended both temporally (Sheehan 1977) and geographically (Jackson and Johnson 2001; Johnson 2003), (3) a common cause having driven both the rock and fossil records simultaneously (Sepkoski 1976; Peters and Foote 2002; Foote 2003; Peters 2005, 2006; Peters and Heim 2010) or (4) redundancy between sampling proxies and apparent diversity (Benton 2010; Benton *et al.* 2011; Dunhill 2012). As accepted by most palaeontologists, it is likely that any particular fossil record reflects a complex mix of rock bias, effort, common cause and redundancy, and it may be hard to distinguish purely 'geological' and 'biological' signals, despite suggestions to the contrary (e.g. Smith 2007; Smith and McGowan 2007; Barrett *et al.* 2009).

Sampling is a combination of geological and human factors (Raup 1972) and can be split into three aspects: (1) sedimentary rock volume, (2) accessibility and (3) worker effort (Raup 1972; Benton *et al.* 2011). The most commonly used sampling proxies are effort by palaeontologists

doi: 10.1111/j.1475-4983.2012.01149.x

(Sheehan 1977; Purnell and Donoghue 2005; Bernard *et al.* 2010), estimates of sedimentary rock volume (Raup 1972, 1976; Kalmar and Currie 2010), rock outcrop (i.e. geological map) area (Smith 2001; Crampton *et al.* 2003; Smith and McGowan 2005, 2007; Uhen and Pyenson 2007; Mander and Twitchett 2008; McGowan and Smith 2008; Marx 2009; Wall *et al.* 2009; Mannion and Upchurch 2010; Peters and Heim 2010), number of sedimentary rock 'packages' bounded by hiatuses (Peters 2005, 2006; Peters and Heim 2010), number or area of fossiliferous localities per time interval (Fara 2002; Barnosky *et al.* 2005; Fountaine *et al.* 2005; Lloyd *et al.* 2008; Butler *et al.* 2011) or number of sedimentary formations per time interval (Peters and Foote 2001, Peters and Foote 2002; Fountaine *et al.* 2005; Wang and Dodson 2006; Fröbisch 2008; Barrett *et al.* 2009; Butler *et al.* 2009, 2011; Benson *et al.* 2010; Marx and Uhen 2010).

The majority of palaeodiversity studies are carried out on a global or continental scale, and the geological sampling proxies used are, at best, vague approximations and do not accurately represent the actual amount of rock available for sampling (Twitchett *et al.* 2000; Crampton *et al.* 2003; Benton 2010; Benton *et al.* 2011; Dunhill 2011, 2012). Very few palaeodiversity studies have been undertaken at small temporal and/or geographical scales. The few that exist have discovered little evidence of sedimentary rock or human sampling bias (e.g. Wignall and Benton 1999; Benton *et al.* 2004; Mander and Twitchett 2008).

The Lower Jurassic of the Dorset Coast has been a mecca for fossil collecting over more than two centuries (Lord *et al.* 2010) and represents the most intensively sampled marine section of this age in the UK (Simms 2004), if not the world. The diverse fossil assemblages are well documented, including 14 type specimens of reptiles (Benton and Spencer 1995), many fishes (Dineley and Metcalf 1999) and a substantial number of invertebrate species. The section also benefits from exceptionally detailed stratigraphic accounts, especially by Lang (1924, 1936; Lang *et al.* 1923, 1928; Lang and Spath 1926) and Buckman (1910, 1917, 1922), who devised high-resolution bed-by-bed stratigraphical frameworks for the Lias (Simms 2004; Lord *et al.* 2010), which have provided the basis for modern lithostratigraphical interpretation.

Given the long history and intensity of study, diverse fossil fauna and detailed bed-by-bed stratigraphic context, the Lower Jurassic of the Dorset Coast represents a perfect test case for investigating sampling bias at a stratigraphically and temporally constrained scale.

GEOLOGICAL SETTING

The Lower Jurassic Series spans about 22 million years from 200 to 178 million years ago (Ma) and comprises the Hettangian to Toarcian stages, which correspond to the lithostratigraphical Lias Group in the UK (Simms 2004) (Fig. 1). The Dorset and East Devon Coast provides the most continuous section of Lower Jurassic exposure in the UK (Simms 2004) and stretches for around 15 km from Pinhay Bay in the west, to Burton Bradstock in the east (Fig. 2), forming part of the Dorset and East Devon Coast World Heritage Site (Lord *et al.* 2010). The Lias Group consists of marine mudstones, sandstones and limestones, most recently revised by Cox *et al.* (1999) into five formations: Blue Lias, Charmouth Mudstone, Dyrham, Beacon Limestone and Bridport Sand (Fig. 1). Many older names have been retained as members (e.g. Eype Clay, Down Cliff Sand, Thorncombe Sand, Marlstone Rock), with the Lower Lias members having been revised recently by Page (2009), with the Shales-with-beef and Black Ven Marls combined into the Black Ven Mudstone, and the Belemnite Marls and Green Ammonite Beds renamed as the Stonebarrow Marl and Seatown Marl respectively (Fig. 1).

The Blue Lias Formation consists of 26–32 m of pale grey limestone interbedded with dark, organic-rich laminated shale and marl (Hallam 1960; Simms 2004; Paul *et al.* 2008). The Charmouth Mudstone Formation is divided into three members (Fig. 1): the Black Ven Mudstone Member (71–79 m) comprises mostly dark mudstones with numerous bands of fibrous calcite ('beef'), together with a few limestone bands; the Stonebarrow Marl Member (23–29 m) consists of blue-grey mudstone alternating with paler marl; the Seatown Marl Member (15–37 m) comprises blue-grey mudstone, with occasional thin limestone beds, becoming increasingly sandy towards the top, except for the top 2 m that are a dark, highly pyritic, mudstone. The Dyrham Formation comprises three members: the Eype Clay Member (67–78 m) consists largely of grey, silty mudstone with a few sandstone layers; the Down Cliff Sand Member (22–28 m) consists mainly of grey-brown sandy mudstone, becomes sandier towards the top and grades into the Thorncombe Sand Member; Thorncombe Sand Member (26–28 m) consists a unit of yellow-weathering sandstones with frequently cemented 'doggers'. The Beacon Limestone Formation (3.2–4.6 m) is divided into the pink conglomeratic limestone of the Marlstone Rock Member and the pale conglomeratic limestone of the Eype Mouth Limestone Member. The Bridport Sand Formation (51–64 m) has in its basal 21 m the Down Cliff Clay Member consisting of silty sandstone, which become sandier upwards.

FOSSIL FAUNA

The diverse fossil assemblages of the Dorset and East Devon Coast Lias are well recorded and consist of a larger

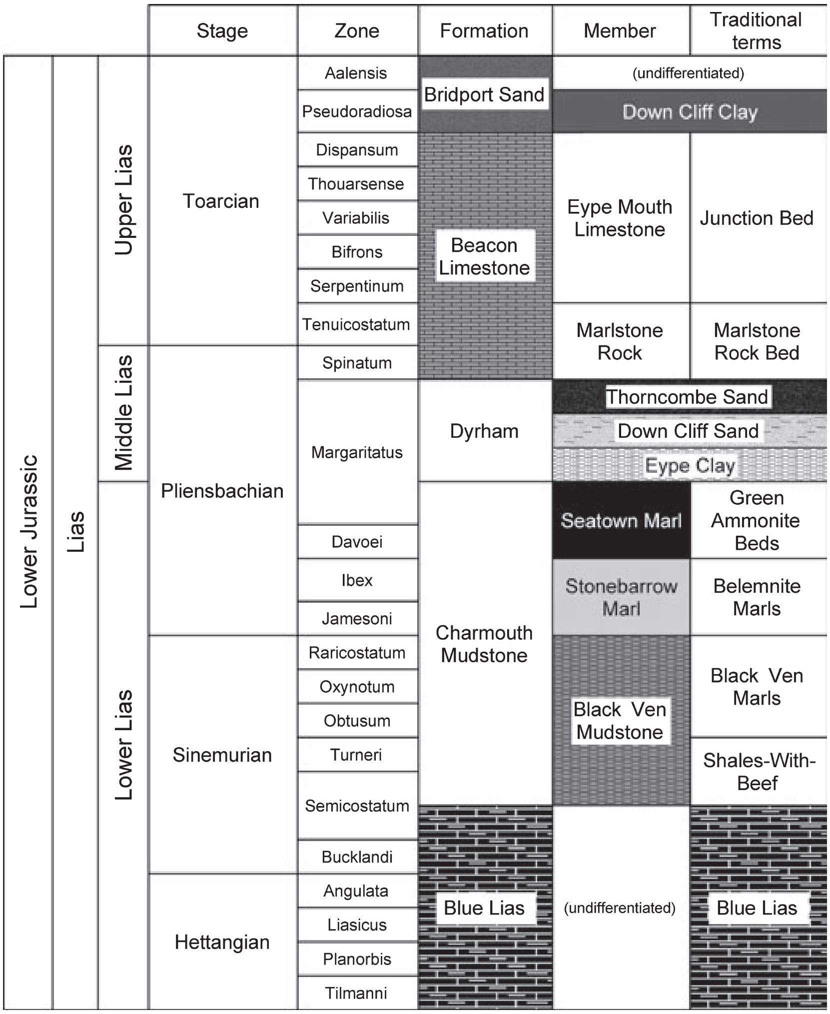


FIG. 1. Chrono-, litho- and biostratigraphy of the Lower Jurassic of the Dorset Coast using revised formational framework of Cox *et al.* (1999) and revised stratigraphic framework of the Lower Lias of Page (2009). Units used in the analysis are identified by lithological symbols that match with Figure 2.

number of type specimens than any other Lower Jurassic section in Britain (Simms 2004). Benthic communities are well represented in the Blue Lias Formation (Lang 1924; Hallam 1960; Simms 2004), which also yields a diverse and abundant ammonite (Hallam 1960), vertebrate (Lord *et al.* 2010) and ichnofossil fauna (Moghadam and Paul 2000). The Black Ven Mudstone is perhaps the most intensively sampled unit and is particularly rich in ammonites (Lang *et al.* 1923; Lang and Spath 1926; Simms 2004), as well as important fossils of marine reptiles (Martill 1995; McGowan and Milner 1999), the dinosaur *Scelidosaurus* (Martill *et al.* 2000) and insects (Zeuner 1962; Whalley 1985). The Stonebarrow Marl displays a sparse benthic assemblage and is dominated by belemnites and ammonites, of which only the belemnites are abundant and well preserved throughout (Lang *et al.* 1928; Simms 2004). Nonbenthic fauna such as ammonites are abundant in the Seatown Marl (Lang 1936; Spath 1936), as are foraminiferans (Macfadyen 1941). The member also yields a diverse benthic assemblage of bivalves, gastropods (Cox 1936) and brachiopods (Muir-Wood 1936). Fossil material is generally scarce in the Eype Clay (Simms 2004), apart from the rich benthic faunas of bivalves, brachiopods, crinoids and crustaceans in the Eype Nodule Bed (Howarth 1957) and Day's Shell Bed (Palmer 1966). The Down Cliff Clay Member is rich in echinoderms, particularly crinoids (Simms 1989, 2004), ophiuroids (Simms 2004) and trace fossils (Sellwood *et al.* 1970). The Thorncombe Sand possesses a diverse, but sporadic, benthic community of brachiopods and gastropods (Simms 2004) as well as a diverse ostracod fauna (Lord 1974). The Beacon Limestone yields a diverse benthic fauna of brachiopods (Ager 1956, 1990), bivalves and gastropods (Jackson 1926), with a rich, nonbenthic ammonite community further up into the Eype Mouth Limestone (Howarth 1992). The Bridport Sand is generally devoid of identifiable fossil material, with only robust belemnite (Doyle 1990–1992) and crinoid (Simms 1989) specimens in much of the Down Cliff Clay, apart from the uppermost beds of the Bridport Sand, which yield abundant ammonites (Simms 2004).

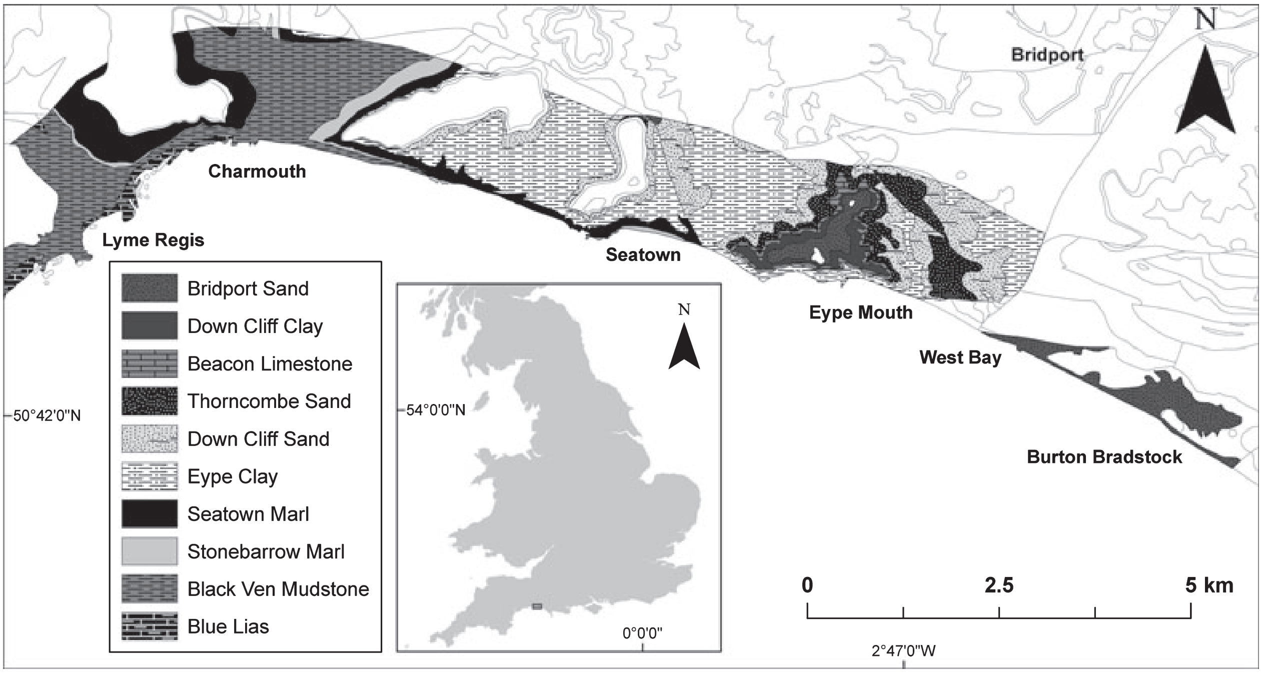


FIG. 2. Map of Lower Jurassic outcrop on the Dorset Coast from Pinhay Bay to Burton Bradstock showing the extent of geological formations and members used in this investigation.

METHODOLOGY

Stratigraphical framework

The new stratigraphy for the Lower Lias, proposed by Page (2009), was used along with the stratigraphy of Cox *et al.* (1999) and Simms (2004) for the Middle and Upper Lias (Fig. 1). However, the Beacon Limestone Formation was not split into its two members, the Marlstone Rock Member and the Eype Mouth Limestone Member (Simms 2004), as the former is only about 15 cm thick, making it impossible to map accurately at a 1:50 000 scale (Fig. 2).

Diversity data

Taxon occurrence data were recorded for all major fossil groups and for each stratigraphic unit under investigation. The records were acquired from a detailed literature search of both monographic publications (e.g. Howarth 1958, 1992; Simms 1989; Doyle 1990–1992) and reviews (e.g. Macfadyen 1941; Barnard 1949; Howarth 1957; Hallam 1960; Zeuner 1962; Palmer 1966; Lord 1974; Whalley 1985) and accounts of individual fossil finds (e.g. Martill 1995; Donovan 1998, 2006; McGowan and Milner 1999). The diversity data from the literature were supplemented by collection records from the British Geological Survey, Keyworth (BGS), Bristol City Museum (BRSMG) and Dorset County Museum, Dorchester (DORCM).

Various problems were encountered while compiling the taxic database. Many of the museum labels and publications are pre-20th century, and as a result, locality information is often vague and taxonomic nomenclature inaccurate. Much of the 19th century material suffers from a lack of accompanying stratigraphical and geographical information, as a detailed stratigraphical framework was not available until the early 20th century. For example, many specimens in the BGS, BRSMG and DORCM collections are labelled simply ‘Lower Lias, Lyme Regis’, which could place the specimen in any of two formations or three members (Blue Lias or Charmouth Mudstone Formation; Black Ven Mudstone, Stonebarrow Marl or Seatown Marls Member). This was resolved by comparing the specimens with stratigraphical occurrences reported in the literature for all of North-West Europe, which proved to be an effective method for groups with restricted ranges, such as ammonoids. However, many genera of other groups considered in this study range through much of the Lower Jurassic, so this proved less satisfactory. The literature search was the best source for generic and species occurrence data, and, of the three collections used, the BGS had the best stratigraphic resolution, with 63.5 per cent of occurrences assigned to one or more of the stratigraphic units used in this study. The BRSMG and DORCM collections are slightly less well constrained, with 56 and 55.4 per cent of taxa, respectively, assigned to one or more of the stratigraphic units used in this study. Taxa that could not be assigned accurately were omitted from the analyses, but almost all of these taxa were, nevertheless, recorded through the literature search.

There is a risk that the museum and literature palaeodiversity curves might show autocorrelation because of the use of literature data to refine stratigraphic ranges of certain museum specimens. This risk is, however, minimized by three factors: first, the literature refinements applied only to dating of ammonoids and not the other taxa; second, the literature was worldwide, or at least focused on the North-west European provinces, not simply the Dorset and East Devon coast; and third, the BGS collections were reported directly from labels with limited correction from the literature, and those uncorrected BGS data correlate well with the corrected museum records (BRSMG, DORCM).

The fossil groups included in the study have been revised taxonomically to differing extents. Some groups have been revised recently, for example, echinoderms (Simms 1989) and belemnites (Doyle 1990–1992, 2003), and the 100 species of reptiles once reported from the Lyme Regis area have since been reduced to 14 (Benton and Spencer 1995). Other groups, however, suffer from nonuniform taxonomic classification and high rates of synonymy. Ammonoids have a potentially overinflated generic richness, and almost certainly a vastly exaggerated species richness (Lord *et al.* 2010), as a result of the historical practices of naming new genera and species on every minor variation in shell morphology, and separate classification of sexually dimorphic macro- and microconches. Ammonoid diversity may well be artificially enhanced by as many as 4–6 times in the absence of recent monographic studies (Lord *et al.* 2010). Taxonomic inflation is also a problem in other fossil groups; for example, the shark genera *Palidiplospinax* and *Palaeospinax*, which occur as the species *Palidiplospinax enniskilleni*, *Palidiplospinax occultidens* and *Palaeospinax priscus* in the Blue Lias and Charmouth Mudstone formations, are all to be revised to a single species of *Palidiplospinax* (S. Klug, pers. comm. 2010).

Sampling proxies

Sampling proxies can be placed into three categories: (1) sedimentary rock volume, (2) accessibility of sedimentary rock exposures and (3) effort by palaeontologists (cf. Raup 1972; Benton *et al.* 2011). Sedimentary rock volume measurements are simple to compile for each stratigraphic unit, with outcrop (geological map) area or unit thickness being suitable metrics from which volume can be estimated.

Accessibility is harder to quantify, and yet must be considered because of the enormously variable patterns of exposure exhibited by the different geological units (Dunhill 2011, 2012). As a first attempt, Dunhill (2011, 2012) used remote sensing and a Geographic Information System (GIS) to record rock exposure area on a regional scale and thus document rock that can actually be sampled. Rock exposure is a better proxy for current accessibility than rock outcrop area, especially when it transpires that many lithostratigraphical units in temperate zones are <10 per cent exposed, with the remainder covered by Pleistocene deposits, soils and human constructions such as cities (Dunhill 2011, 2012). Further, it should be noted that metrics of both outcrop and exposure area fail to address the problem of large, lithologically homogeneous exposures that may yield the same fossil assemblages over vast areas and through long times of study. The number of fossiliferous localities, as used by Fara (2002), Benton *et al.* (2004) and Lloyd *et al.* (2008), may be a better proxy. The number of fossiliferous localities is easy to measure when compiling data on inland 'islands' of rock exposure (i.e. quarries and road/rail cuttings), but is difficult to quantify on coastal sections where exposure is continuous. Accessibility can also be tested by considering the distance of exposures from points of access and centres of population (e.g. car parks, public houses, shops).

A key issue about accessibility is the historical dimension. In many areas of the world, access to certain geological formations has varied from high to low levels, as quarries and other industrial excavations have permitted extensive collecting, often for a short time, and then no collecting at all after these exposures are lost. However, in our case study here, the East Devon and Dorset coastline has offered more or less continuous exposure of all sampled geological formations over the centuries. The short-term availability of quarries and road cuts may then have had only a modest confounding effect. There are data sets on commercial quarries in the UK, together with their years of operation, and rail and road cuttings can be dated approximately, so in future work, we propose to explore such historical influences on collecting.

Research effort by palaeontologists is also difficult to quantify. Ideally, one should use data on the amount of sediment sampled per stratigraphic unit or the number of field days spent studying each section. Although useful when carrying out small-scale field sampling projects (e.g. Mander *et al.* 2008), this method is obviously impractical when considering all the palaeontological sampling of the Lower Jurassic of the Dorset and East Devon Coast over the past two centuries. Sheehan (1977) devised the Paleontological Interest Unit (PIU), an estimate of the number of marine invertebrate palaeontologists studying different time periods of the Phanerozoic. He found that the PIUs correlated positively with Raup's (1972, 1976) estimates of Phanerozoic diversity and preserved sedimentary rock volumes; he concluded that the amount of sedimentary rock preserved influences systematic effort, which in turn drives apparent diversity. However, Raup (1977) was quick to point out that 'systematists follow the fossils'; in other words, palaeodiversity probably drives PIUs, because palaeontologists are likely to work most on rock sections that yield abundant fossils and avoid rock sections known to yield little of interest.

In this investigation, several proxies were used to represent each of the three aspects of sampling. *Rock volume* was calculated using average stratigraphic unit thickness (in metres) obtained from several sources (Cope *et al.* 1980; House 1993; Simms 2004) and rock outcrop (geological map) area for each stratigraphic unit (obtained from BGS digital bedrock geology DiGMapgb-50 of the UK (1 : 50 000) in ArcGIS 9.3).

Accessibility could not be based on numbers of fossiliferous localities along the Dorset Coast because of the continuous exposure in cliff faces, so other metrics were required. The methodology of Dunhill (2011, 2012) was used to calculate rock exposure area as a measure of accessibility, using remote sensing in the form of Google Earth imagery and ArcGIS 9.3, for each stratigraphic unit. ArcGIS 9.3 was also used to record the number of car parks and public houses within 100 m, 500 m and 1 km of exposures for each stratigraphic unit.

Effort is also hard to quantify. It is impossible to obtain data on the number of field days spent by many palaeontologists over the centuries in sampling or the amount of sedimentary rock sampled for each stratigraphic unit of the Lower Jurassic of the Dorset and East Devon Coast. Metrics of worker effort included the number of publications, the number of publications that arose from studies involving field sampling and the number of authors who have published on each stratigraphic unit. The obvious limitation of these metrics is the lack of a quantifiable measure for the effort of individual commercial and non-commercial fossil collectors who have undoubtedly expended significant effort collecting in areas such as Lyme Regis and Charmouth, but who may not provide material for museums and publications in scientific journals.

Topographical models

As most of the rock exposure on the Dorset Coast occurs in cliff sections, it is necessary to take topography into account when making areal measurements. Therefore, triangular irregular networks (TINs) were created, in ArcGIS 9.3, from NextMap Britain 5 m Digital Terrain Model (DTM) grids (vertical accuracy ±0.6 m) to allow the calculation of true surface area including topographical irregularity (a three-dimensional measure), rather than just using map surface area (a two-dimensional measure). The TINs were used to calculate 3D outcrop area, volume and exposure area for each stratigraphic unit. The 3D values are of course consistently higher than the 2D measures because of the addition of topographical complexity.

Testing for bias in the fossil record

Correlations were sought between generic and species diversity in the whole taxonomic data set as well as for individual fossil groups. This is to assess whether these two taxonomic levels show the same pattern and to help identify any effects of synonymy in the species data. Correlations were also sought between diversity of the different fossil groups to detect any similarity in diversity patterns between groups. Diversity similarities between groups might suggest sampling-driven diversity or a real pattern of diversity, while differences in diversity patterns between groups might suggest the influence of changing palaeoenvironments preferred by different groups or the differential effects of preservational or sampling biases. Strong correlations between the sampling proxies and diversity would suggest that the fossil record of the Lower Jurassic of the Dorset Coast is biased by sampling. Benjamini & Hochberg (1995) false discovery rate corrections were applied to control for the large number of correlation tests.

RESULTS

Diversity

Both generic and species diversity follow the same pattern (Fig. 3), with strong positive correlations, even after correction for multiple comparisons, between the two taxonomic levels in the entire faunal data set and in the diversity curves of all the major fossil groups (Table 1). The only exception is the belemnites, which show a greatly inflated species count in the Stonebarrow Marl Member; the correlation between species and generic diversity, although significant, is not so strong and becomes nonsignificant after correction for multiple comparisons (Table 1).

The diversity data from the three collections and the literature search all follow the same trend (Fig. 4; Table 2), suggesting that they all capture the true pattern of apparent diversity with reasonable accuracy. The three museum collections sampled are all widely regarded as comprehensive collections from the Dorset coast, and the fact that they all agree, both with each other and with the overall literature-based palaeodiversity signal, suggests that further synopses of museum data are unlikely to change the conclusion significantly.

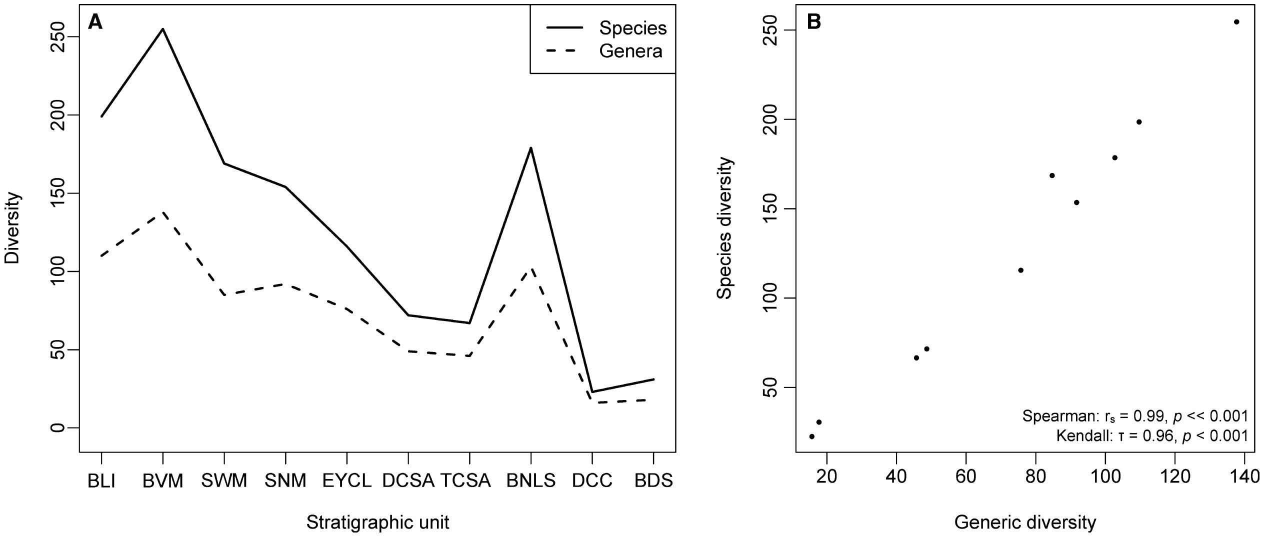


FIG. 3. A, Total generic and species diversity through the Lower Jurassic of the Dorset Coast. BNLS, Beacon Limestone Formation; BLI, Blue Lias Formation; BDS, Bridport Sand Formation; BVM, Black Ven Mudstone Member; DCC, Down Cliff Clay Member; DCSA, Down Cliff Sand Member; EYCL, Eype Clay Member; SWM, Stonebarrow Marl Member; SNM, Seatown Marl Member; TCSA, Thorncombe Sand Member. B, Correlation between total generic and total species diversity.

TABLE 1. Generic and species diversity correlation coefficients.

	Spearman rank		Kendall tau	
	r_s	p	τ	p
All fossil groups	0.99	<<0.001**	0.96	<0.001**
Microfossils	0.98	<0.001**	0.95	<0.001**
Brachiopods	0.98	<0.001**	0.95	<0.001**
Bivalves	0.98	<0.001**	0.93	<0.001**
Gastropods	0.88	<0.001**	0.79	0.002**
Cephalopods	0.96	<0.001**	0.89	<0.001**
Ammonites	0.98	<0.001**	0.93	<0.001**
Belemnites	0.65	0.04*	0.57	0.03*
Echinoderms	0.98	<0.001**	0.95	<0.001**
Vertebrates	1	<<0.001**	1	<0.001**

Statistically significant correlations ($p < 0.05$) before correction for multiple comparisons indicated by asterisks (*) or double asterisks (**) for statistically significant correlations after correction for multiple comparisons using the false discovery rate procedure of Benjamini & Hochberg (1995).

Total diversity is highest through the Lower Lias, particularly in the Black Ven Mudstone Member, before falling through the Middle Lias, followed by a sharp rise in the Beacon Limestone Formation, before a final fall to its lowest level through the rest of the Upper Lias (Fig. 3). Brachiopod, bivalve, all cephalopod, ammonoid, echinoderm and vertebrate diversity correlate well with, and therefore follow a similar pattern to, total diversity (Fig. 5; Tables S1, S2). However, many of these correlations break down when applying the correction for multiple comparisons, with the exception of all cephalopods, the ammonoids and the echinoderms (Fig. 5; Tables S1, S2). Microfossil, gastropod and belemnite diversity curves differ significantly from the total diversity curve (Fig. 5; Tables S1, S2). Microfossils are most diverse during the time of the Blue Lias Formation and Seatown Marl Member (Fig. 5), and gastropods most diverse through the Middle Lias and sparse in the Lower and Upper Lias (Fig. 5). Belemnites are most diverse in the Stonebarrow Marl Member and Beacon Limestone Formation, and relatively sparse in all other parts of the sequence (Fig. 5).

Sampling proxies

The proxies for sedimentary rock volume (average thickness, outcrop area and estimated volume) correlate positively with one another (Figs 6A–C, 7A; Table S3), with highest rock volumes in the Black Ven Mudstone and Eype Clay members, and lowest rock volumes in the Blue Lias and Beacon Limestone formations and Down Cliff Clay and Stonebarrow Marl members (Fig. 6A–C; Table S3). However, after the application of the correction for multiple comparisons, average thickness no longer correlates significantly with the other proxies for sedimentary rock volume (Table S3). All proxies for accessibility (rock exposure, number of car parks within 1 km and number of public houses within 1 km) correlate positively with each other, but the correlations become nonsignificant after correction for multiple comparisons (Figs 6D, F–G; Table S3). The Seatown Marl

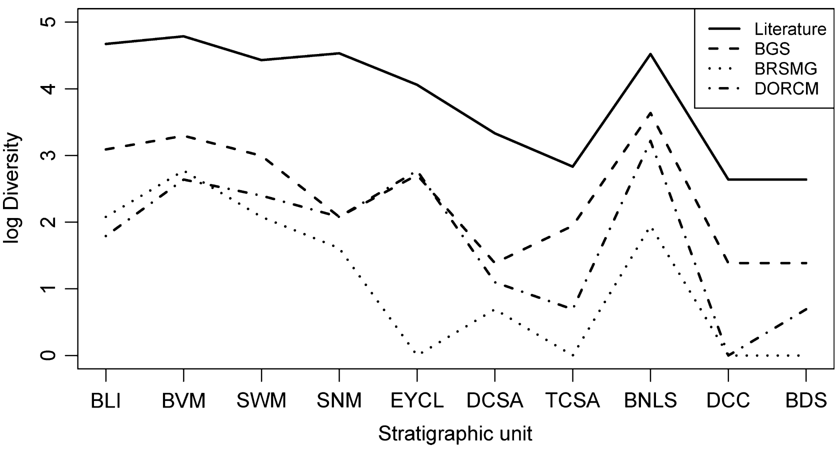


FIG. 4. Generic diversity recorded from a literature search and collection records from the British Geological Survey (BGS), Bristol City Museum (BRSMG) and Dorset County Museum (DORCM).

TABLE 2. Fossil collection correlation coefficients.

	BGS	BRSMG	DORCM
Literature search	r_s = 0.83, p = 0.003** τ = 0.7, p = 0.005**	r_s = 0.88, p < 0.001** τ = 0.72, p = 0.004**	r_s = 0.69, p = 0.03* τ = 0.52, p = 0.04*
BGS	NA	r_s = 0.77, p = 0.01* τ = 0.76, p = 0.02*	r_s = 0.83, p = 0.003** τ = 0.67, p = 0.008**
BRSMG	NA	NA	r_s = 0.63, p = 0.05* τ = 0.54, p = 0.04*

Statistically significant correlations ($p < 0.05$) before correction for multiple comparisons indicated by asterisks (*) or double asterisks (**) for statistically significant correlations after correction for multiple comparisons using the false discovery rate procedure of Benjamini & Hochberg (1995).

Member is the most accessible unit, showing large areas of exposure and close proximity to both car parks and public houses (Fig. 6D, F–G). The Black Ven Mudstone and Eype Clay members are also very accessible, while the Thorncombe Sand and Down Cliff Clay members, along with the Beacon Limestone Formation, are least accessible (Fig. 6D, F–G). The proxies representing worker effort correlate positively until the application of correction for multiple comparisons, when the correlations between the number of publications involving field sampling and the number of authors per stratigraphic unit become nonsignificant (Fig. 6E; Table S3). Worker effort is mostly directed towards the Lower Lias, with the Blue Lias Formation and the Black Ven Mudstone Member being associated with the highest number of publications and authors (Fig. 6E). Both the Middle and Upper Lias show far fewer publications and authors than the units of the Lower Lias (Fig. 6E).

None of the rock volume proxies correlates with any of the sampling proxies representing accessibility or worker effort (Figs 6–7; Table S3), apart from a positive correlation between average thickness and the number of car parks within 1 km of exposed rock, which subsequently becomes nonsignificant once correction for multiple comparisons has been applied (Table S3). Rock exposure, the number car parks and the number of public houses correlate with the number of publications and the number of publications with actual sampling (Figs 6–7; Table S3); however, these correlations also become nonsignificant after correction for multiple comparisons.

Diversity and sampling proxies

There are no significant correlations between total generic or species diversity and any of the proxies representing rock volume (thickness, outcrop and estimated volume) (Figs 8A–C, 9A–E; Tables S4, S5). There are significant correlations between total diversity and rock exposure, but only when considering the 2D data (Figs 8D, 9F–G; Tables S4, S5), and this correlation becomes nonsignificant when corrected for multiple comparisons (Tables S4, S5). There are significant correlations between total diversity and the number of public houses within 1 km, although the correlation between total diversity and the number of car parks within 1 km is not significant (Figs 8F, 9K–L;

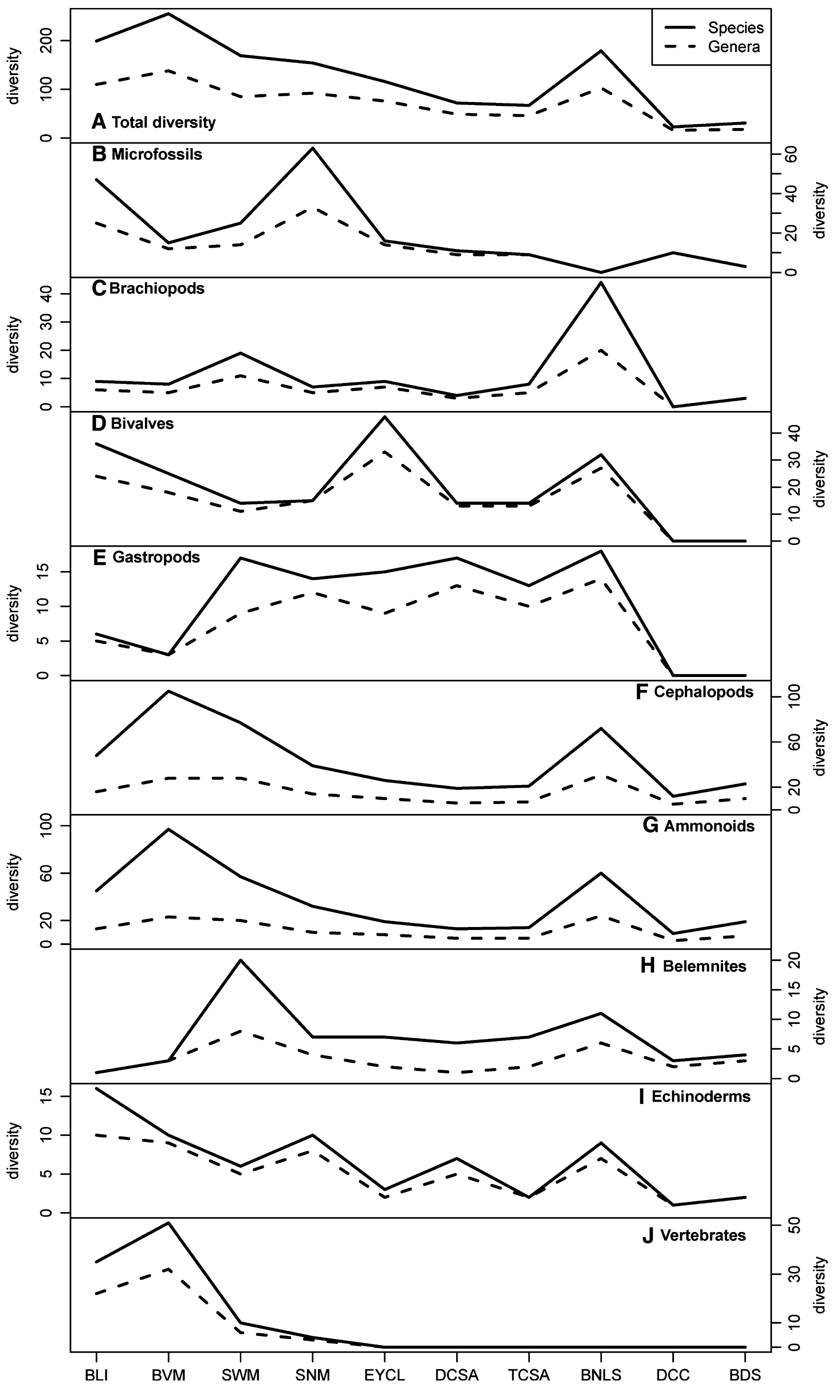


FIG. 5. Generic and species diversity of: A, all fossil groups; B, microfossils; C, brachiopods; D, bivalves; E, gastropods; F, Cephalopods; G, ammonoids; H, belemnites; I, echinoderms; J, vertebrates.

Tables S4, S5). There are strong correlations between total diversity and the number of publications, the number of publications with sampling and the number of authors having worked on each stratigraphic unit (Figs 8E, 9H–J, Tables S4, S5).

As with the total diversity, there are no correlations between any of the proxies for rock volume and the diversity of individual fossil groups (Tables S4, S5). Gastropod, belemnite and brachiopod diversity do not correlate with any of the sampling proxies (Tables S4, S5). Bivalve generic diversity does not correlate with any sampling proxy (Tables S4, S5), although bivalve species diversity does correlate with rock exposure (2D) (Fig. 10D, E) and the number of publications with sampling (Table S5), yetboth correlations become nonsignificant when corrected for multiple comparisons (Table S5). Microfossil generic and species diversity correlates well with rock exposure area (2D and 3D) (Fig. 10A) as well

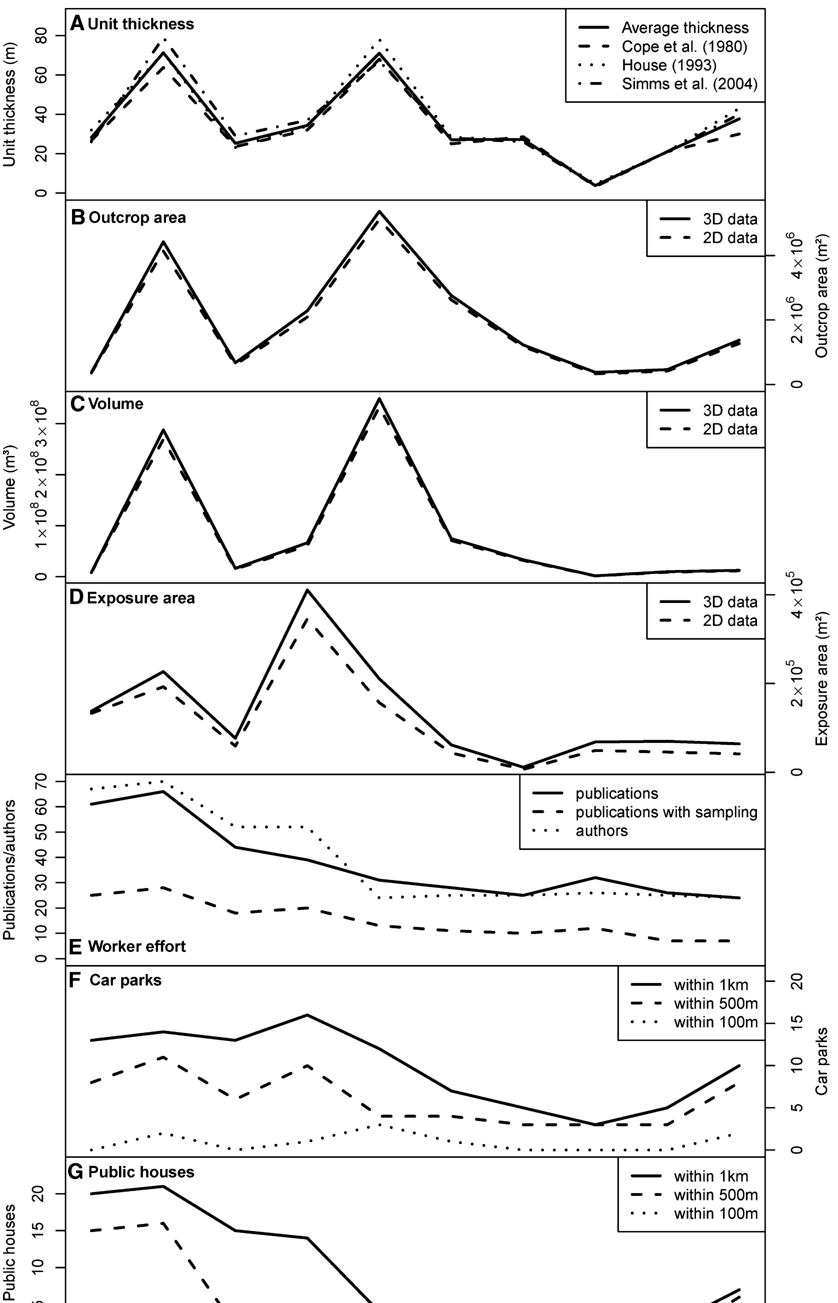


FIG. 6. Sampling proxies: A, unit thickness; B, outcrop area; C, estimated volume; D, exposure area; E, worker effort; F, number of car parks within distance of exposed localities; G, number of public houses within distance of exposed localities.

as with the number of publications, with sampling and with car parks and public houses within 1 km of exposed sedimentary rock (Fig. 10B; Tables S4, S5); however, only the correlation with car parks remains significant after correction for multiple comparisons (Tables S4, S5). All cephalopod, ammonoid, echinoderm and vertebrate generic and species diversities correlate well with the number of publications, the number of publications with sampling and the number of authors (Fig. 10F–J; Tables S4, S5). However, only the correlations involving echinoderms and vertebrates remain significant after correction for multiple comparisons (Tables S4, S5). All cephalopod, ammonoid, echinoderm and vertebrate generic diversities also correlate well with the number of public houses within 1 km of exposed sedimentary rock (Fig. 10K; Tables S4, S5), and while echinoderm species diversity does not significantly correlate with the number of public houses within 1 km, all cephalopod, ammonoid and

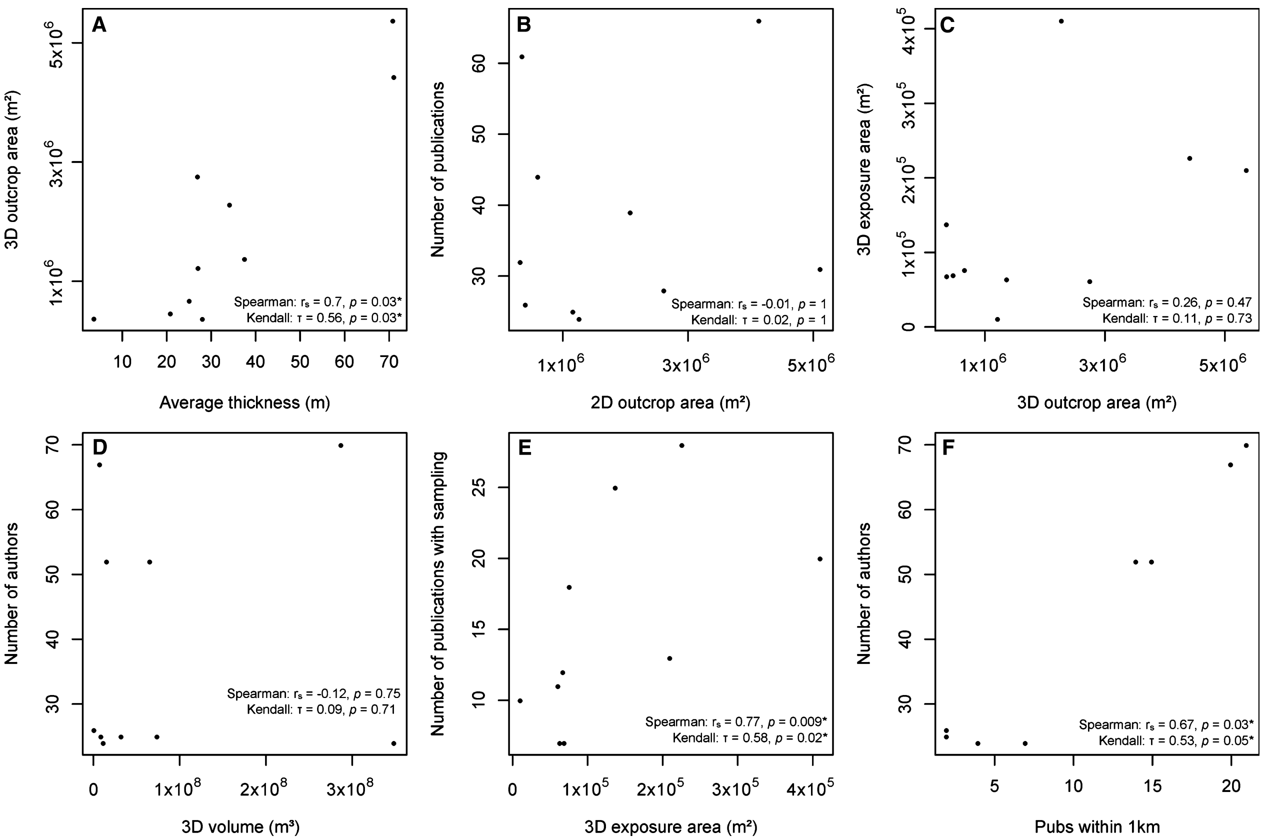


FIG. 7. Correlation plots of sampling proxies: A, average thickness and 3D outcrop area; B, 2D outcrop area and number of publications per stratigraphic unit; C, 3D outcrop and 3D exposure area; D, 3D estimated volume and number of authors having worked on each stratigraphic unit; E, 3D exposure area and number of publications with sampling per stratigraphic unit; F, number of public houses within 1 km of exposed localities and number of authors having worked on each stratigraphic unit. Statistically significant correlations ($p < 0.05$) before correction for multiple comparisons indicated by asterisks (*) or double asterisks (**) for statistically significant correlations after correction for multiple comparisons using the false discovery rate procedure of Benjamini and Hochberg (1995).

vertebrate species diversities do (Table S5), but of all these correlations, only that for vertebrates remains significant after correcting for multiple comparison (Tables S4, S5). Vertebrate generic diversity also correlates significantly with the number of car parks within 1 km and the amount of rock exposure, in 2D and 3D (Fig. 10L; Table S4), and echinoderm species diversity correlates with rock exposure area (2D) (Table S5). However, only the correlation between vertebrates and car parks within 1 km remains significant after correction for multiple comparisons (Table S4).

DISCUSSION

Diversity

Palaeodiversity is not consistent throughout the Lower Jurassic of the Dorset Coast. The similar patterns followed by, and the strong correlations between, generic and species diversity in all groups combined, and in all major fossil group, suggest that the taxonomic level used in the analysis does not significantly alter the results, a finding and recommendation of previous studies (Sepkoski and Kendrick 1993; Robeck *et al.* 2000; Smith 2007). However, belemnite generic and species diversity shows a weaker correlation that becomes nonsignificant after correction for multiple comparisons. Belemnite species richness is particularly high in the Stonebarrow Marl Member, formerly known as the ‘Belemnite Marls’ because of the abundance of belemnite fossils (Doyle, *in* Lord *et al.* 2010) (Fig. 1). Perhaps the species count is overinflated, and synonymy may be a problem among Lower Lias belemnites, and although Doyle (2003) identified many synonyms, true species diversity may still be lower in the Stonebarrow Marl than indicated here. The ammonoids show a high species-to-genus ratio throughout the entire Lower Jurassic, particularly during the

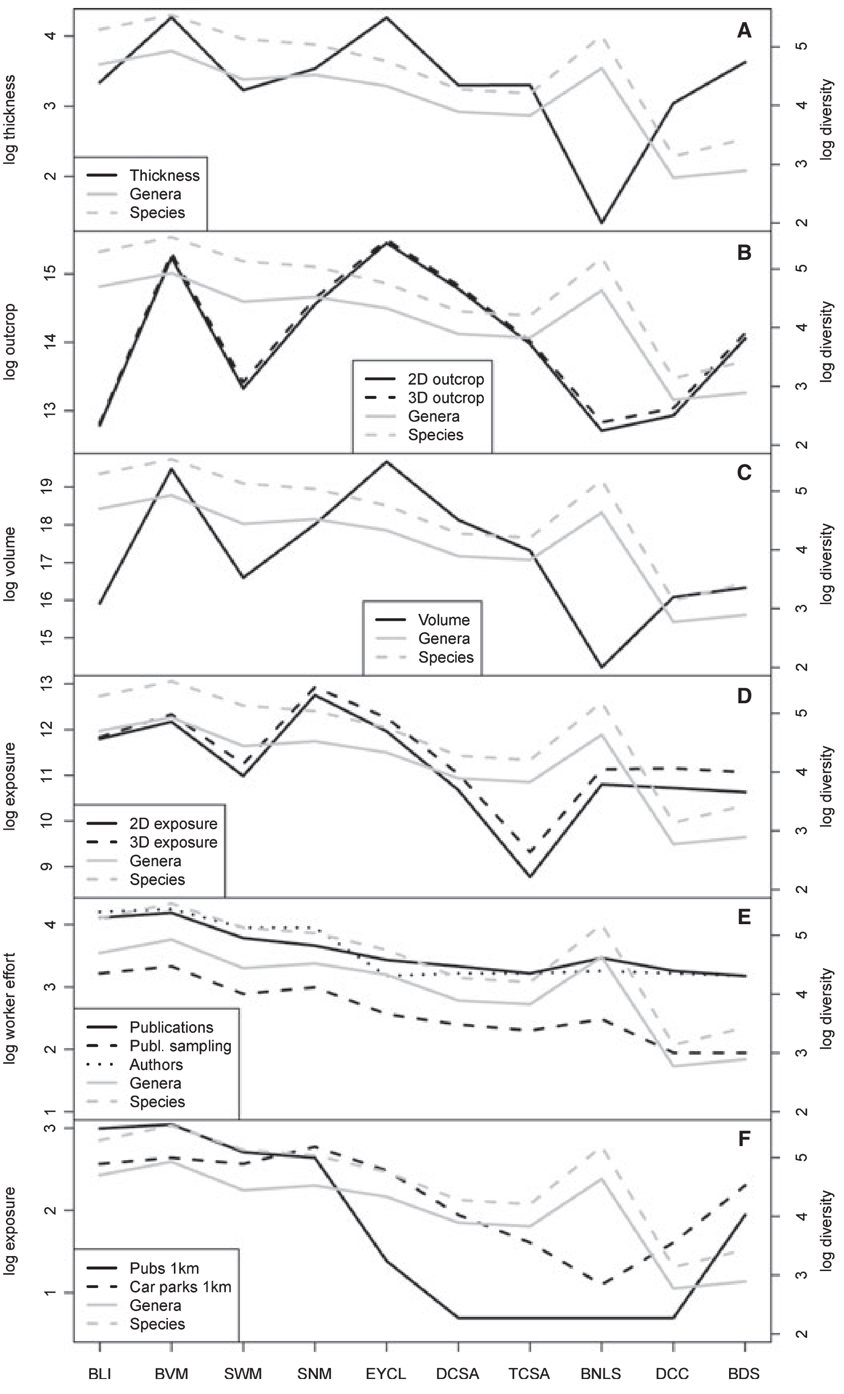
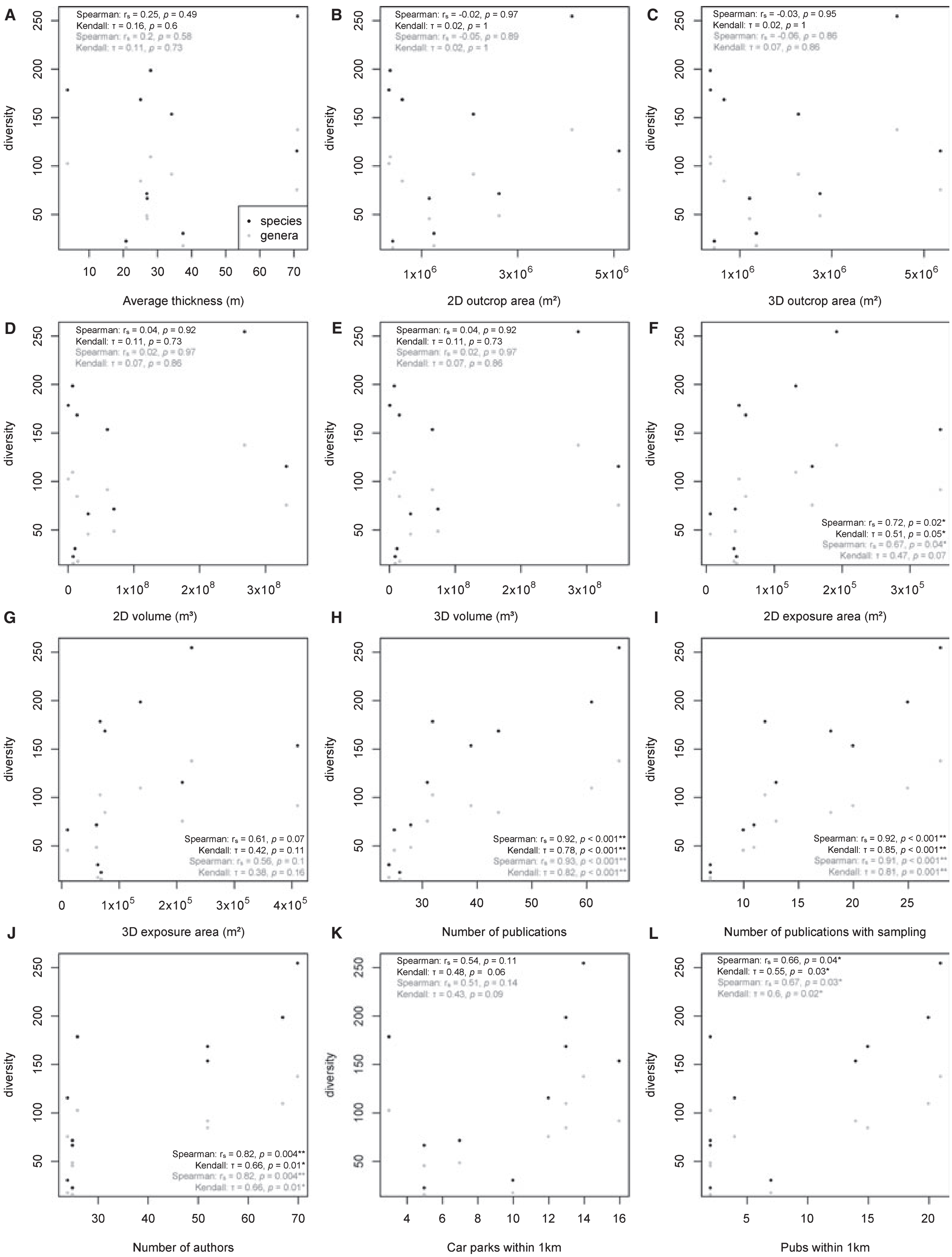


FIG. 8. Generic and species diversity plotted against sampling proxies through the Lower Jurassic of the Dorset Coast. A, average unit thickness; B outcrop area; C, estimated volume; D, exposure area; E, worker effort; F, access and proximity to amenities.

FIG. 9. Correlation plots of total diversity and sampling proxies. A, average thickness; B, 2D outcrop area; C, 3D outcrop area; D, 2D estimated volume; E, 3D estimated volume; F, 2D exposure area; G, 3D exposure area; H, number of publications per stratigraphic unit; I, number of publications with sampling per stratigraphic unit; J, number of authors having worked on each stratigraphic unit; K, number of car parks within 1 km of exposed localities; L, number of public houses within 1 km of exposed localities. Statistically significant correlations ($p < 0.05$) before correction for multiple comparisons indicated by asterisks (*) or double asterisks (**) for statistically significant correlations after correction for multiple comparisons using the false discovery rate procedure of Benjamini & Hochberg (1995).


A
diversity
Spearman: r_s = 0.25, p = 0.49
Kendall: τ = 0.16, p = 0.6
Spearman: r_s = 0.2, p = 0.58
Kendall: τ = 0.11, p = 0.73
species
genera
Average thickness (m)
B
Spearman: r_s = -0.02, p = 0.97
Kendall: τ = 0.02, p = 1
Spearman: r_s = -0.05, p = 0.89
Kendall: τ = 0.02, p = 1
2D outcrop area (m²)
C
Spearman: r_s = -0.03, p = 0.95
Kendall: τ = 0.02, p = 1
Spearman: r_s = -0.06, p = 0.86
Kendall: τ = 0.07, p = 0.86
3D outcrop area (m²)
D
Spearman: r_s = 0.04, p = 0.92
Kendall: τ = 0.11, p = 0.73
Spearman: r_s = 0.02, p = 0.97
Kendall: τ = 0.07, p = 0.86
2D volume (m³)
E
Spearman: r_s = 0.04, p = 0.92
Kendall: τ = 0.11, p = 0.73
Spearman: r_s = 0.02, p = 0.97
Kendall: τ = 0.07, p = 0.86
3D volume (m³)
F
Spearman: r_s = 0.72, p = 0.02*
Kendall: τ = 0.51, p = 0.05*
Spearman: r_s = 0.67, p = 0.04*
Kendall: τ = 0.47, p = 0.07
2D exposure area (m²)
G
Spearman: r_s = 0.61, p = 0.07
Kendall: τ = 0.42, p = 0.11
Spearman: r_s = 0.56, p = 0.1
Kendall: τ = 0.38, p = 0.16
3D exposure area (m²)
H
Spearman: r_s = 0.92, p < 0.001**
Kendall: τ = 0.78, p < 0.001**
Spearman: r_s = 0.93, p < 0.001**
Kendall: τ = 0.82, p < 0.001**
Number of publications
I
Spearman: r_s = 0.92, p < 0.001**
Kendall: τ = 0.85, p < 0.001**
Spearman: r_s = 0.91, p < 0.001**
Kendall: τ = 0.81, p = 0.001**
Number of publications with sampling
J
Spearman: r_s = 0.82, p = 0.004**
Kendall: τ = 0.66, p = 0.01*
Spearman: r_s = 0.82, p = 0.004**
Kendall: τ = 0.66, p = 0.01*
Number of authors
K
Spearman: r_s = 0.54, p = 0.11
Kendall: τ = 0.48, p = 0.06
Spearman: r_s = 0.51, p = 0.14
Kendall: τ = 0.43, p = 0.09
Car parks within 1km
L
Spearman: r_s = 0.66, p = 0.04*
Kendall: τ = 0.55, p = 0.03*
Spearman: r_s = 0.67, p = 0.03*
Kendall: τ = 0.6, p = 0.02*
Pubs within 1km

Lower Lias. While it is likely that the species diversity is significantly inflated by over splitting (Page, *in* Lord *et al.* 2010), the species diversity curve correlates extremely well with the generic diversity curve, suggesting that synonymy does not significantly alter the overall diversity trend through the system.

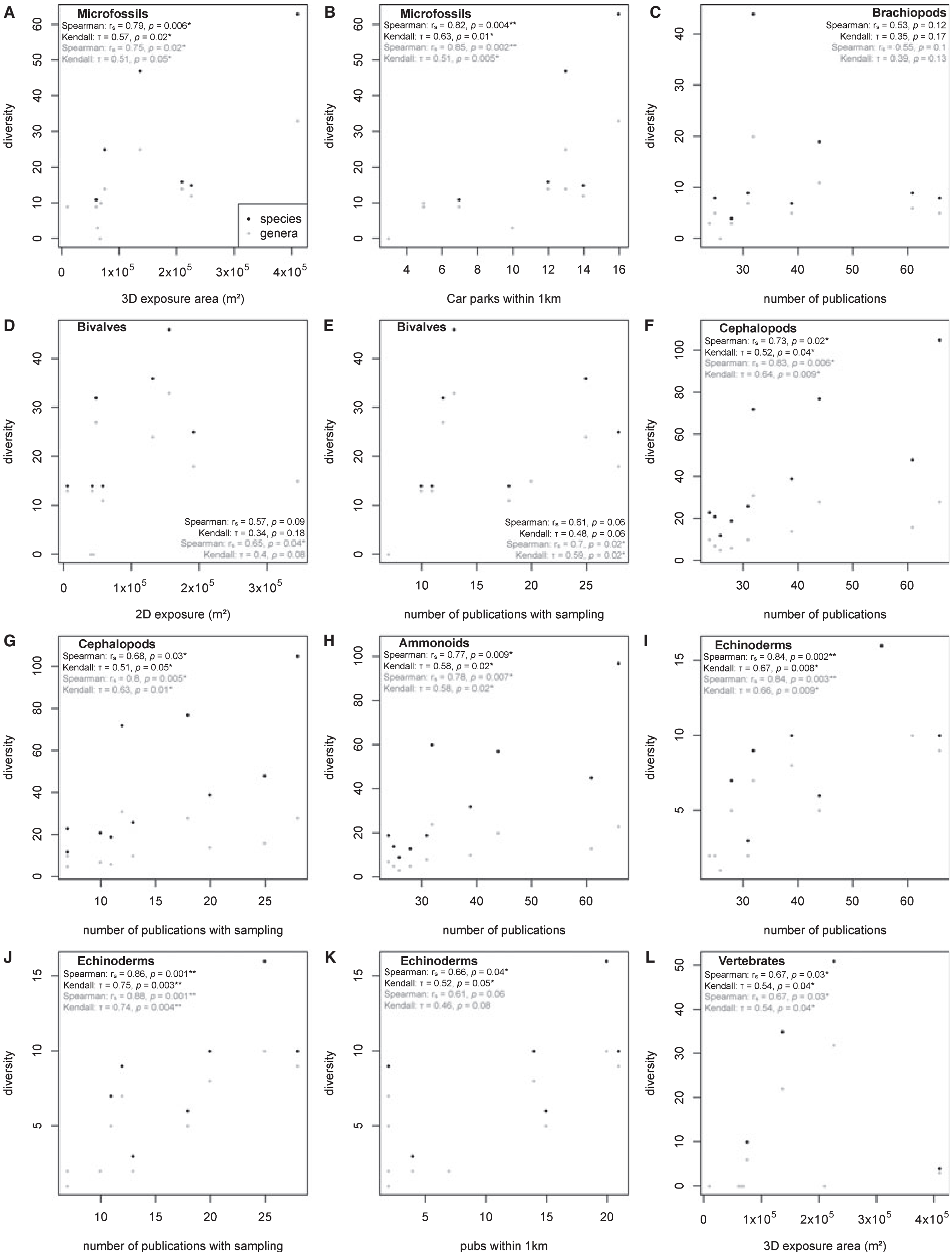
The diversity patterns of different fossil groups are largely similar to the pattern of total diversity, with only microfossils, gastropods and belemnites recording markedly different diversity trajectories. However, the diversity curves of different fossil groups peak during different stratigraphic units; for example, diversities of ammonoids and vertebrates peak during the Black Ven Mudstone Member, brachiopods during the Beacon Limestone Formation, bivalves during the Eype Clay Member and echinoderms during the Blue Lias Formation. This could be because (1) apparent diversity represents a true picture of biological diversity, with microfossils, gastropods and belemnites displaying significantly different diversity trajectories than all the other major groups; (2) different stratigraphic units represent different facies that supported higher or lower diversities of different faunal groups; (3) stratigraphic units differ in their preservation potential of different groups; or (4) sampling biases influenced different fossil groups to varying degrees. The vertebrates provide an interesting example, in that they have recorded occurrences only in the Lower Lias (Blue Lias Formation to Seatown Marl Member), with no identifiable specimens from the Middle and Upper Lias. There are four possible explanations for this gap in the record: (1) this is a true picture of biological diversity more widely, and vertebrates were entirely absent; (2) the facies represent environments in which vertebrates were rare; (3) vertebrate bone was initially buried in the sediments, but subsequently lost during diagenesis; or (4) collectors have failed to find specimens that are actually present. The first is unlikely because marine reptiles are known from the Middle and Upper Lias elsewhere (e.g. Holzmaden, Ilminster, Yorkshire), and the lineages continue across this time interval in cladograms. The third suggestion requires investigation, but there are no traces of 'ghosts' of dissolved bones, and the sedimentology of all parts of the Lias is comparable. The fourth suggestion is unlikely because all units are well exposed and accessible, and there are strong incentives for collectors to find such specimens. There are perhaps local reasons, whether facies or geographical barriers that kept marine reptiles from the Dorset area during the time of deposition of the Middle and Upper Lias.

Sampling proxies

As average thickness, outcrop area (from the 2D and 3D data) and volume (from 2D and 3D data) all correlate positively, all three proxies presumably track the same phenomenon and are all similar representations of the amount of Lower Jurassic sedimentary rock on the Dorset and East Devon Coast. However, this does not mean that all this rock is accessible and therefore available for sampling. As with other areas of England and Wales (Dunhill 2011), rock exposure area (from the 2D or 3D data) on the Dorset and East Devon Coast does not correlate with any of the proxies representing rock volume, suggesting that average thickness, outcrop area and estimated volume are not good proxies for the amount of currently accessible sedimentary rock. The rock volume proxies also show no correlation with any of the worker effort proxies, suggesting that the number of authors working, or the number of publications, on a particular stratigraphic unit is not governed by overall geographical extent. Palaeontologists presumably, therefore, search according to the bonanza principle (where am I most likely to find my prize?) rather than working systematically over all of the exposure, even if most of the time little is found. This bears out Raup's (1977) insight that effort and diversity may correlate because palaeontologists go where they know fossils are abundant and diverse.

The total number of publications, and the number of publications with actual field sampling, per stratigraphic unit correlates well with rock exposure (from the 2D and 3D data) suggesting that the amount of worker effort may be influenced by the amount of sedimentary rock that is accessible and available for sampling. However, the correlations between rock exposure area and the number of authors who worked on each stratigraphic unit are weaker. This could suggest either that the various metrics of effort are themselves measuring different things or that palaeontologists indeed go where they know fossils occur

FIG. 10. Correlation plots of individual fossil group diversity and sampling proxies. A, 3D exposure and microfossils; B, car parks within 1 km and microfossils; C, number of publications and brachiopods; D, 2D exposure area and bivalves; E, number of publications with sampling and bivalves; F, number of publications and cephalopods; G, number of publications with sampling and cephalopods; H, number of publications and ammonoids; I, number of publications and echinoderms; J, number of publications with sampling and echinoderms; K, public houses within 1 km and echinoderms; L, 3D exposure area and vertebrates. Statistically significant correlations ($p < 0.05$) before correction for multiple comparisons indicated by asterisks (*) or double asterisks (**) for statistically significant correlations after correction for multiple comparisons using the false discovery rate procedure of Benjamini & Hochberg (1995).


A
Microfossils
Spearman: rs = 0.79, p = 0.006*
Kendall: τ = 0.57, p = 0.02*
Spearman: rs = 0.75, p = 0.02*
Kendall: τ = 0.51, p = 0.05*
species
genera
diversity
3D exposure area (m²)
B
Microfossils
Spearman: rs = 0.82, p = 0.004**
Kendall: τ = 0.63, p = 0.01*
Spearman: rs = 0.85, p = 0.002**
Kendall: τ = 0.51, p = 0.005*
Car parks within 1km
C
Brachiopods
Spearman: rs = 0.53, p = 0.12
Kendall: τ = 0.35, p = 0.17
Spearman: rs = 0.55, p = 0.1
Kendall: τ = 0.39, p = 0.13
number of publications
D
Bivalves
Spearman: rs = 0.57, p = 0.09
Kendall: τ = 0.34, p = 0.18
Spearman: rs = 0.65, p = 0.04*
Kendall: τ = 0.4, p = 0.08
2D exposure (m²)
E
Bivalves
Spearman: rs = 0.61, p = 0.06
Kendall: τ = 0.48, p = 0.06
Spearman: rs = 0.7, p = 0.02*
Kendall: τ = 0.59, p = 0.02*
number of publications with sampling
F
Cephalopods
Spearman: rs = 0.73, p = 0.02*
Kendall: τ = 0.52, p = 0.04*
Spearman: rs = 0.83, p = 0.006*
Kendall: τ = 0.64, p = 0.009*
number of publications
G
Cephalopods
Spearman: rs = 0.68, p = 0.03*
Kendall: τ = 0.51, p = 0.05*
Spearman: rs = 0.8, p = 0.005*
Kendall: τ = 0.63, p = 0.01*
number of publications with sampling
H
Ammonoids
Spearman: rs = 0.77, p = 0.009*
Kendall: τ = 0.58, p = 0.02*
Spearman: rs = 0.78, p = 0.007*
Kendall: τ = 0.58, p = 0.02*
number of publications
I
Echinoderms
Spearman: rs = 0.84, p = 0.002**
Kendall: τ = 0.67, p = 0.008*
Spearman: rs = 0.84, p = 0.003**
Kendall: τ = 0.66, p = 0.009*
number of publications
J
Echinoderms
Spearman: rs = 0.86, p = 0.001**
Kendall: τ = 0.75, p = 0.003**
Spearman: rs = 0.88, p = 0.001**
Kendall: τ = 0.74, p = 0.004**
number of publications with sampling
K
Echinoderms
Spearman: rs = 0.66, p = 0.04*
Kendall: τ = 0.52, p = 0.05*
Spearman: rs = 0.61, p = 0.06
Kendall: τ = 0.46, p = 0.08
pubs within 1km
L
Vertebrates
Spearman: rs = 0.67, p = 0.03*
Kendall: τ = 0.54, p = 0.04*
Spearman: rs = 0.67, p = 0.03*
Kendall: τ = 0.54, p = 0.04*
3D exposure area (m²)

(Raup 1977), and yet then the discrepancy between publication counts and author counts must be explained. This might arise because the publication count indeed covaries with exposure, indicating some association of the two measures, but the number of workers does not, perhaps because of dominance at different times by a small number of highly prolific authors (e.g. Buckman, Lang). Perhaps the 'number of authors' metric used in previous studies (e.g. Sheehan 1977) is not useful except in cases where overall numbers of authors are consistently high, so mixing prolific and nonprolific publishers; otherwise, as here, a single author can turn out 20 papers a year, or none at all.

It is evident that the sampling proxies devised to cover the three aspects of sampling (sedimentary rock volume, accessibility and worker effort) do not covary. Therefore, the many studies that have used single proxies, such as estimates of sedimentary rock outcrop area (Smith 2001; Smith and McGowan 2005, 2007; Uhen and Pyenson 2007; Marx 2009; Wall *et al.* 2009), volume (Raup 1972, 1976; Kalmar and Currie 2010) or worker effort (Sheehan 1977), are only addressing specific aspects of sampling. The lack of correlation between sampling proxies means that is unlikely that a single sampling metric will be found to 'correct' an empirical fossil occurrence data set (Dunhill 2012), and it makes sense to use many proxies, reflecting the different aspects of sampling, to test for bias at outcrop in the fossil record.

Diversity and sampling proxies

Many recent studies have discovered a correlation between palaeodiversity and proxies for sedimentary rock volume on a global (Raup 1972, 1976; Wall *et al.* 2009), continental (Smith 2001; Smith and McGowan 2005, 2007) and national scale (Crampton *et al.* 2003). At a local scale, however, there appears to be no evidence for covariance between proxies for rock volume and palaeodiversity, at least in well-studied regions such as this one. Through the Lower Jurassic of the Dorset and East Devon Coast, neither unit thickness, rock outcrop area or estimated volume correlate with total diversity or diversity of any of the major fossil groups, suggesting that sedimentary rock volume is not a major influence on apparent diversity in this test case.

The correlation observed between rock exposure area and total diversity, as well as between rock exposure area and microfossil, vertebrate, bivalve and echinoderm diversity, suggests that the accessibility of sedimentary rock may have influenced apparent diversity. This claim is strengthened by observed correlations between the number of public houses within 1 km of exposed localities and total diversity, as well as between the number of pubs and microfossil, cephalopod, ammonoid, echinoderm and vertebrate diversity, in addition to correlations between the number of car parks within 1 km of exposed localities and microfossil and vertebrate diversity, suggesting that proximity to both roads and centres of population may have an influence on sampling regimes. Equally, it is likely that villages, farms, roads and pubs are built in areas that are not topographically challenging and so share those aspects of accessibility with the most-visited geological localities. However, while microfossil and vertebrate diversity correlate with exposure areas calculated from both 2D and 3D data, total diversity only correlates well with rock exposure area calculated from the 2D data and not so when accounting for topographical changes. As the majority of rock exposure on the Dorset Coast occurs in cliff sections, the 3D data represent a more accurate approximation of exposure area. Therefore, if the accessibility of sedimentary rock was controlling apparent diversity, it would be expected that total diversity would correlate well with rock exposure calculated from the 3D data. However, it may be that not all exposed rock is actually available for sampling, as steep cliff exposures are not easily or safely accessible. It must also be acknowledged that the correlations between the accessibility proxies and diversity are not particularly robust, and only a small majority of the initial significant positive correlations remain so after correction for multiple comparisons (Tables S4, S5).

Historical changes in rock availability might also play a major role. For example, along the Dorset and Devon Coast, geomorphology has changed from time to time as a result of landslips (Denness *et al.* 1975; Brunsden and Jones 1976; Pitts 1983; Brunsden and Moore 1999). These events generate a sudden and rich patch of exposure, but only for a short time, before the rock and debris are grassed over or washed away by the sea. Construction projects also facilitate short-term access, as for example, when a railway line, road or major building is constructed and excavations expose rock temporarily. Local interest has also doubtless been important; it is well known that enthusiastic amateur or professional palaeontologists may at times hunt for fossils daily or weekly in particular localities, and then, there will be times when nobody visits a section for years on end. The ever-changing nature of the eroding Dorset coastline and varying collecting effort over the centuries make it difficult to measure rock accessibility over long time periods. We shall explore these historical aspects in future studies.

There is a strong positive correlation between total diversity and both the number of publications and the number of publications with sampling. The significance of correlations between both the number of publications and the number of publications with sampling and all the individual fossil groups is consistently strong, with the

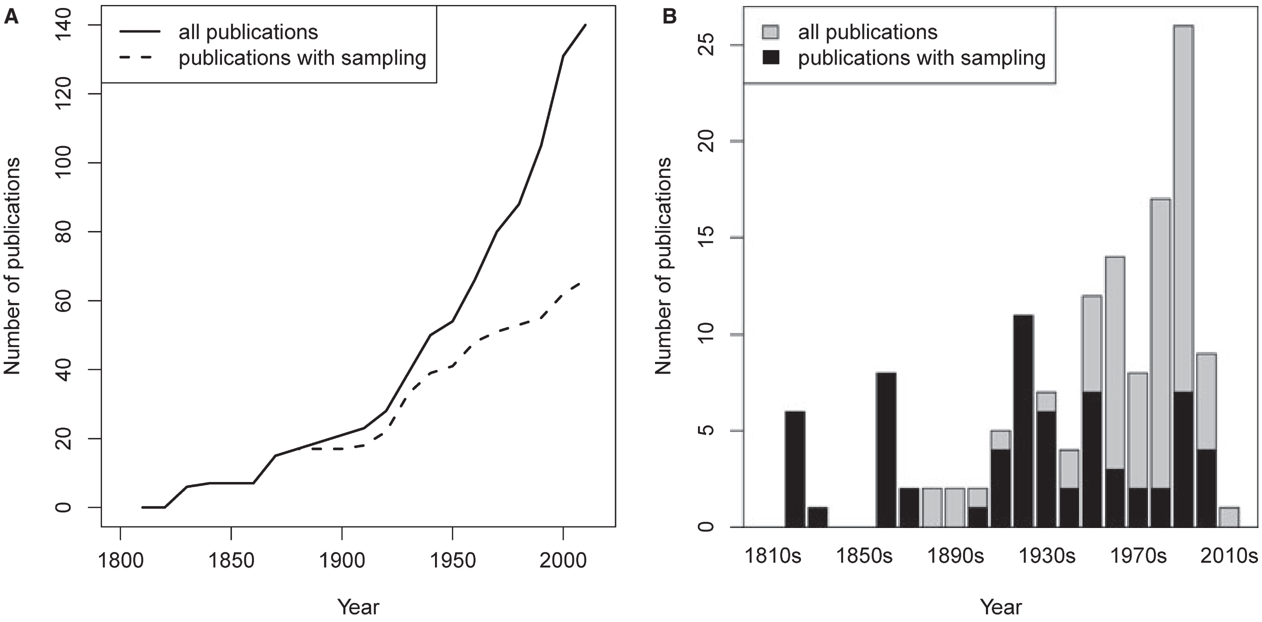


FIG. 11. A, Cumulative number of publications on the Lower Jurassic of the Dorset Coast 1800–2010 and B, Total number of publications on the Lower Jurassic of the Dorset Coast per decade 1800–2010.

exception of brachiopods, gastropods and belemnites, and many of the these correlations remain highly significant even after correction for multiple comparisons. Worker effort is, therefore, likely to have had a significant influence on apparent diversity (Sheehan 1977). This claim is strengthened by positive correlations between the number of authors and total diversity, in addition to correlations between the number of authors and cephalopod, ammonoid, echinoderm and vertebrate diversity, of which many remain highly significant after correction for multiple comparisons. However, these correlations could just as easily be interpreted the other way around, with diversity driving the number of publications and the number of authors working on particular stratigraphic units (Raup 1977; Purnell and Donoghue 2005), as noted earlier. In the case of the Lower Jurassic of the Dorset and Devon Coast, the latter is more likely, as this series of sections are one of the most heavily sampled stratigraphic systems in the world, and there has never been a shortage of collectors on the ground (Simms 2004; Lord *et al.* 2010). However, publication intensity seems to have increased through the 20th century, peaking in the 1990s, before falling after 2000 (Fig. 11). This would suggest that sampling intensity on the Dorset Coast is still high, but the majority of recent publications are reviews or revisions of existing material. The number of publications involving actual field sampling peaked in the 1920s, before falling steadily through the latter half of the 20th century (Fig. 11B). This could reflect a change in work practices over the last 100 years, with decreasing emphasis on prolonged and detailed field collecting or that palaeontologists believe that nearly every species in the Dorset Lias has been found and identified, and further intensive searching would yield little new.

CONCLUSIONS

This is the first case study to investigate bias in the fossil record of an intensely sampled division of geological time in a restricted area. We have considered a number of sampling proxies used in earlier work to provide insights into fossil record bias in terms of rock volume, accessibility and effort. Importantly, we have added some new sampling proxies that benefit from innovative GIS techniques, namely map-based (2D) exposure, topographically corrected (3D) exposure, proximity to roads and buildings, and unit volume. Our most startling conclusion is that the sampling proxies representing rock volume, accessibility and effort do not correlate positively with each other, and so none can be regarded as better or worse than the others, nor can any be regarded as a definitive sampling proxy that encompasses all elements of sampling bias. These results support the use of the equal-grid sampling approach of Smith and McGowan (2007), where the recording of fossil bearing rocks assumes effective sampling in an area. However, the Lower Jurassic outcrop of the Dorset and Devon coast is an exceptional case of thorough sampling, and these assumptions are not likely to hold true for less intensively sampled outcrops. The fact that species and generic diversity in the Dorset and East Devon Lias do not correlate with traditional sampling proxies for rock volume, such as stratigraphic unit thickness, outcrop area or rock volume, suggests the need for caution in the wider and uncritical use of such measures to explore large-scale data sets (e.g. Smith 2007). However, this study does not offer an insight into the effect of adding more sampling regions and thus increasing beta (= regional)-diversity as a result of wider sampling of facies. The covariation between palaeodiversity and measures of effort (publication counts, worker counts) is a useful discovery that probably indicates a bonanza effect, namely that palaeontologists are drawn to collect in proportion to the known richness of fossils in different formations and locations.

If supposed sampling proxies such as these metrics of human effort are at least partially redundant with the palaeodiversity signal, as is suggested by the plausible two-way feeding between the signals (fossils are rich in unit A because it has been biased by over-collection vs many fossils have been found in unit A because life was truly diverse at that time and many palaeontologists have been drawn to study them), then care should be exercised in dubbing one covarying signal (e.g. number of publications, exposure area) as the sampling metric and the other (palaeodiversity) as the biased signal that is in need of correction (Benton *et al.* 2011). Once historical factors of changing accessibility have been further explored, further light may be shed on the role of sampling on our perception of palaeodiversity.

Acknowledgements. Thanks to Paul Shepherd (British Geological Survey), Roger Vaughan (Bristol City Museum), Jenny Cripps and Paul Tomlinson (Dorset County Museum) for the provision of museum specimens and data. Thanks to Kevin Page, Chris Paul, John Callomon and Stefanie Klug for advice and discussion on various fossil groups. This work was funded by NERC doctoral training grant NE/H525111/1 and BGS CASE grant to AMD, and NERC grant NE/C518973/1 to MJB. We also thank Alistair McGowan and Andrew Smith for reading the manuscript and providing many helpful comments.

Editor. Patrick Orr

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table S1. Correlation coefficients between generic diversity of different fossil groups.

Table S2. Correlation coefficients between species diversity of different groups.

Table S3. Correlation coefficients between sampling proxies.

Table S4. Correlation coefficients between generic diversity and sampling proxies.

Table S5. Correlation coefficients between species diversity and sampling proxies.

REFERENCES

AGER, D. V. 1956. Some new Liassic Terebratuloids. *Proceedings of the Geologists Association*, **67**, 1–14.

—— 1990. *British Liassic Terebratulida (Brachiopoda): part 1*. Monograph of the Palaeontographical Society, Palaeontographical Society, London.

BARNARD, T. 1949. Foraminifera from the Lower Lias of the Dorset coast. *Quarterly Journal of the Geological Society*, **105**, 347–388.

BARNOSKY, A. D., CARRASCO, M. A. and DAVIS, E. B. 2005. The impact of the species-area relationship on estimates of palaeodiversity. *PLoS Biology*, **3**, e266.

BARRETT, P. M., MCGOWAN, A. J. and PAGE, V. 2009. Dinosaur diversity and the rock record. *Proceedings of the Royal Society of London B*, **276**, 2667–2674.

BENJAMINI, Y. and HOCHBERG, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical, Series B*, **57**, 289–300.

BENSON, R. B. J., BUTLER, R. J., LINDGREN, J. and SMITH, A. S. 2010. Mesozoic marine tetrapod diversity: mass extinctions and temporal heterogeneity in geological megabiases affecting vertebrates. *Proceedings of the Royal Society of London B*, **277**, 829–834.

BENTON, M. J. 2010. The origins of modern biodiversity on land. *Philosophical Tranactions of the Royal Society of London B*, **365**, 3667–3679.

—— and SPENCER, P. S. 1995. *Fossil reptiles of Great Britain*. The Geological Conservation Review Series, Joint Nature Conservation Committee, Chapman & Hall, Peterborough, 386 pp.

—— TVERDOKHLEBOV, V. P. and SURKOV, M. V. 2004. Ecosystem re-modelling among vertebrates at the Permian-Triassic boundary in Russia. *Nature*, **432**, 97–100.

—— DUNHILL, A. M., LLOYD, G. T. and MARX, F. G. 2011. Assessing the quality of the fossil record: insights from vertebrates. 63–94. *In* MCGOWAN, A. J. and SMITH, A. B. (eds). *Comparing the geological and fossil records: implications for biodiversity studies. Geological Society, London, Special Publications*, **358**, 63–94 pp.

BERNARD, E. L., RUTA, M., TARVER, J. E. and BENTON, M. J. 2010. The fossil record of early tetrapods: Worker effort and the end-Permian mass extinction. *Acta Palaeontologica Polonica*, **55**, 229–239.

BRUNSDEN, D. and JONES, D. K. C. 1976. The evolution of landslide slopes in Dorset. *Philosophical Tranactions of the Royal Society of London A*, **283**, 605–631.

—— and MOORE, R. 1999. Engineering geomorphology on the coast: lessons from west Dorset. *Geomorphology*, **31**, 391–409.

BUCKMAN, S. S. 1910. Certain Jurassic (Lias-Oolite) strata of south Dorset. *Quarterly Journal of the Geological Society*, **66**, 52–89.

—— 1917. Jurassic chronology: I – Lias. *Quarterly Journal of the Geological Society*, **292**, 257–334.

—— 1922. Jurassic chronology: II-preliminary studies. Certain Jurassic strata near Eypesmouth (Dorset); the Junction-Bed of Watton Cliff and associated rocks. *Quarterly Journal of the Geological Society*, **78**, 378–457.

BUTLER, R. J., BARRETT, P. M., NOWBATH, S. and UPCHURCH, P. 2009. Estimating the effects of the rock record on pterosaur diversity patterns: implications for hypotheses of bird/pterosaur competitive replacement. *Paleobiology*, **35**, 432–446.

—— BENSON, R. B. J., CARRANO, M. T., MANNION, P. D. and UPCHURCH, P. 2011. Sea level, dinosaur diversity and sampling biases: investigating the 'common cause' hypothesis in the terrestrial realm. *Proceedings of the Royal Society of London B*, **278**, 1165–1170.

COPE, J. C. W., GETTY, T. A., HOWARTH, M. K., MORTON, N. and TORRENS, H. S. 1980. *A correlation of the Jurassic rocks in the British Isles. Part one: introduction and Lower Jurassic*. Special Report No. 14, Geological Society, London, 73 pp.

COX, L. R. 1936. The Gastropoda and Lamellibranchia of the Green Ammonite Beds of Dorset. *Quarterly Journal of the Geological Society*, **96**, 456–471.

COX, B. M., SUMBLER, M. G. and IVIMEY-COOK, H. C. 1999. A formational framework for the Lower Jurassic of England and Wales (onshore area). British Geological Survey Research Report, RR/99/01.

CRAMPTON, J. S., BEU, A. G., COOPER, R. A., JONES, C. A., MARSHALL, B. and MAXWELL, P. A. 2003. Estimating the rock volume bias in paleobiodiversity studies. *Science*, **301**, 358–360.

DENNESS, B., CONWAY, B. W., MCCANN, D. M. and GRAINGER, P. 1975. Investigation of a coastal landslip at Charmouth, Dorset. *Quarterly Journal of Engineering Geology and Hydrogeology*, **8**, 119–140.

DINELEY, D. L. and METCALF, S. J. 1999. *Fossil fishes of Great Britain*. The Geological Conservation Review Series. Joint Nature Conservation Committee, Peterborough, 675 pp.

DONOVAN, D. T. 1998. A new ammonite genus from the Lower Jurassic (Upper Sinemurian) of Dorset, England. *Palaeontology*, **41**, 993–999.

—— 2006. Phragmoteuthida (Cephalopoda: Coleoidea) from the Lower Jurassic of Dorset, England. *Palaeontology*, **3**, 673–684.

DOYLE, P. 1990–1992. *The British Toarcian (Lower Jurassic) Belemnites*. Monograph of the Palaeontographical Society, Palaeontographical Society, London, 1–79.

—— 2003. Type belemnites of Simpson's *Fossils of the Yorkshire Lias*. *Proceedings of the Yorkshire Geological Society*, **54**, 147–184.

DUNHILL, A. M. 2011. Using remote sensing and a GIS to quantify rock exposure area in England and Wales: implications for paleodiversity studies. *Geology*, **39**, 111–114.

—— 2012. Problems with using rock outcrop area as a paleontological sampling proxy: rock outcrop and exposure area compared to coastal proximity, topography, land use and lithology. *Paleobiology*, **38**, 840–857.

FARA, E. 2002. Sea-level variations and the quality of the continental fossil record. *Journal of the Geological Society, London*, **159**, 489–491.

FOOTE, M. 2003. Origination and extinction through the Phanerozoic: a new approach. *The Journal of Geology*, **111**, 125–148.

FOUNTAINE, M. R., BENTON, M. J., DYKE, G. J. and NUDDS, R. L. 2005. The quality of the fossil record of Mesozoic birds. *Proceedings of the Royal Society of London B*, **272**, 289–294.

FRÖBISCH, J. 2008. Global taxonomic diversity of Anomodonts (Tetropoda, Therapsida) and the terrestrial rock record across the Permian–Triassic boundary. *PLoS One*, **3**, e3733.

HALLAM, A. 1960. A sedimentary and faunal study of the Blue Lias of Dorset and Glamorgan. *Philosophical Transactions of the Royal Society of London B*, **243**, 1–44.

HOUSE, M. 1993. *Geology of the Dorset Coast*. Geologists Association Guide, London, 162 pp.

HOWARTH, M. K. 1957. The Middle Lias of the Dorset Coast. *Quarterly Journal of the Geological Society*, **112**, 185–204.

—— 1958. *A monograph of the ammonites of the Liassic family Amaltheidae in Britain*. Monograph of the Palaeontographical Society, Palaeontographical Society, London, 1–53.

—— 1992. *The ammonite family Hildoceratidae in the Lower Jurassic of Britain*. Monograph of the Palaeontographical Society, Palaeontographical Society, London, 1–200.

JACKSON, J. F. 1926. The Junction-Bed of the Middle and Upper Lias on the Dorset Coast. *Quarterly Journal of the Geological Society*, **82**, 490–525.

JACKSON, J. B. C. and JOHNSON, K. G. 2001. Measuring past biodiversity. *Science*, **293**, 2401–2404.

JOHNSON, K. G. 2003. New data for old questions. *Paleobiology*, **29**, 19–21.

KALMAR, A. and CURRIE, D. J. 2010. The completeness of the continental fossil record and its impact on patterns of diversification. *Paleobiology*, **36**, 51–60.

LANG, W. D. 1924. The Blue Lias of the Devon and Dorset coasts. *Proceedings of the Geologists Association*, **35**, 169–185.

—— 1936. The Green Ammonite Beds of the Dorset Lias. *Quarterly Journal of the Geological Society*, **92**, 423–437.

—— and SPATH, L. F. 1926. The Black Marl of Black Ven and Stonebarrow, in the Lias of the Dorset Coast. *Quarterly Journal of the Geological Society*, **82**, 144–165.

—— —— and RICHARDSON, W. A. 1923. Shales-with-'Beef,' a sequence in the Lower Lias of the Dorset Coast. *Quarterly Journal of the Geological Society*, **79**, 47–66.

—— —— COX, L. R. and MUIR-WOOD, H. M. 1928. The Belemnite Marls of Charmouth, a series in the Lias of the Dorset Coast. *Quarterly Journal of the Geological Society*, **84**, 179–222.

LLOYD, G. T., DAVIS, K. E., PISANI, D., TARVER, J. E., RUTA, M., SAKAMOTO, M., HONE, D. W. E., JENNINGS, R. and BENTON, M. J. 2008. Dinosaurs and the Cretaceous terrestrial revolution. *Proceedings of the Royal Society of London B*, **275**, 2483–2490.

LORD, A. 1974. Ostracods from the Domerian and Toarcian of England. *Palaeontology*, **17**, 599–622.

LORD, A. R., DAVIS, P. G., DUFFIN, C. J., EDMONDS, R., PAGE, K. N., JOHNSON, B., WHITTAKER, J. E., PAUL, C. R. C., UNDERWOOD, C. J., OWEN, E. F., CLEMENTS, D., PALMER, C. P., TODD, J. A., MUNT, M. C., DOYLE, P., ROSS, A., COLLINS, J. S. H., EWIN, T. A. M., FRASER, N., HUNTER, A. W., LEWIS, D. N., FOREY, P., LONGBOTTOM, A. E., MULLEY, J., MILNER, A. C., WALSH, S. and CHAPMAN, S. D. 2010. *Fossils from the Lower Lias of the Dorset Coast*. The Palaeontological Association, Field Guides to Fossils: Number 13, London, 436 pp.

MACFADYEN, W. A. 1941. Foraminifera from the Green Ammonite Beds, Lower Lias, of Dorset. *Philosophical Transactions of the Royal Society of London B*, **231**, 1–73.

MANDER, L. and TWITCHETT, R. J. 2008. Quality of the Triassic-Jurassic bivalve fossil record in Northwest Europe. *Palaeontology*, **51**, 1213–1223.

———— and BENTON, M. J. 2008. Palaeoecology of the Late Triassic extinction event in the SW UK. *Journal of the Geological Society, London*, **165**, 319–332.

MANNION, P. D. and UPCHURCH, P. 2010. Completeness metrics and the quality of the sauropodomorph fossil record through geological and historical time. *Paleobiology*, **36**, 283–302.

MARTILL, D. M. 1995. An ichthyosaur with preserved soft tissue from the Sinemurian of southern England. *Palaeontology*, **38**, 897–903.

——BATTEN, D. J. and LOYDELL, D. K. 2000. A new specimens of the thyreophoran dinosaur cf. *Scelidosaurus with soft tissue preservation. Palaeontology*, **43**, 549–559.

MARX, F. G. 2009. Marine mammals through time: when less is more in studying palaeodiversity. *Proceedings of the Royal Society of London B*, **276**, 887–892.

—— and UHEN, M. D. 2010. Climate, critters, and cetaceans: Cenozoic drivers of the evolution of modern whales. *Science*, **327**, 993–996.

MCGOWAN, C. and MILNER, A. C. 1999. A new Pliensbachian ichthyosaur from Dorset, England. *Palaeontology*, **42**, 761–768.

MCGOWAN, A. J. and SMITH, A. B. 2008. Are global Phanerozoic marine diversity curves truly global? A study of the relationship between regional rock records and global Phanerozoic marine diversity. *Paleobiology*, **34**, 80–103.

MOGHADAM, H. V. and PAUL, C. R. C. 2000. Trace fossils of the Jurassic, Blue Lias, Lyme Regis, Southern England. *Ichnos*, **7**, 283–306.

MUIR-WOOD, H. M. 1936. Brachiopoda from the Lower Lias, Green Ammonite Beds, of Dorset. *Quarterly Journal of the Geological Society*, **92**, 472–487.

PAGE, K. N. 2009. High resolution ammonite stratigraphy of the Charmouth Mudstone Formation (Lower Jurassic: Sinemurian-Lower Pliensbachian) in south-west England, UK. *Volumina Jurassica*, **7**, 19–29.

PALMER, C. P. 1966. The fauna of Day's Shell Bed in the Middle Lias of the Dorset Coast. *Proceedings of the Dorset Natural History and Archaeological Society*, **87**, 69–80.

PAUL, C. R. C., ALLISON, P. A. and BRETT, C. E. 2008. The occurrence and preservation of ammonites in the Blue Lias Formation (lower Jurassic) of Devon and Dorset, England and their palaeoecological, sedimentological and diagenetic significance. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **270**, 258–272.

PETERS, S. E. 2005. Geologic constraints on the macroevolutionary history of marine animals. *Proceedings of the National Academy of Sciences*, **102**, 12326–12331.

—— 2006. Genus extinction, origination, and the durations of sedimentary hiatuses. *Paleobiology*, **32**, 387–407.

—— and FOOTE, M. 2001. Biodiversity in the Phanerozoic: a reinterpretation. *Paleobiology*, **27**, 583–601.

————2002. Determinants of extinction in the fossil record. *Nature*, **416**, 420–424.

—— and HEIM, N. A. 2010. The geological completeness of paleontological sampling in North America. *Paleobiology*, **36**, 61–79.

PITTS, J. 1983. The temporal and spatial development of landslides in the Axmouth-Lyme Regis Undercliffs National Nature Reserve, Devon. *Earth Surface Processes and Landforms*, **8**, 589–603.

PURNELL, M. A. and DONOGHUE, M. J. 2005. Between death and data: biases in the interpretation of the fossil record of conodonts. *Special Papers in Palaeontology*, **73**, 7–25.

RAUP, D. M. 1972. Taxonomic diversity during the Phanerozoic. *Science*, **177**, 1065–1071.

—— 1976. Species diversity in the Phanerozoic: an interpretation. *Paleobiology*, **2**, 289–297.

—— 1977. Systematists follow the fossils. *Paleobiology*, **3**, 328–329.

ROBECK, H. E., MALEY, C. C. and DONOGHUE, M. J. 2000. Taxonomy and temporal diversity patterns. *Paleobiology*, **26**, 171–187.

SELLWOOD, B. W., DURKIN, M. K. and KENNEDY, W. J. 1970. Field meeting on the Jurassic and Cretaceous rocks of Wessex. *Proceedings of the Geologists Association*, **81**, 715–732.

SEPKOSKI, J. J. Jr 1976. Species diversity in the Phanerozoic: species-area effects. *Paleobiology*, **2**, 298–303.

—— and KENDRICK, D. C. 1993. Numerical experiments with model monophyletic and paraphyletic taxa. *Paleobiology*, **19**, 168–184.

SHEEHAN, P. M. 1977. A reflection of labor by systematics? *Paleobiology*, **3**, 325–328.

SIMMS, M. J. 1989. *British Lower Jurassic Crinoids*. Monograph of the British Palaeontographical Society, Palaeontographical Society, London, 1–103.

——2004. The Wessex Basin (Dorset and central Somerset). 53–107. *In* SIMMS, M. J., CHIDLAW, N., MORTON, N. and PAGE, K. N. (eds). *British Lower Jurassic stratigraphy*. The Geological Conservation Review Series, Joint Nature Conservation Committee, Peterborough, 458 pp.

SMITH, A. B. 2001. Large-scale heterogeneity of the fossil record: implications for Phanerozoic biodiversity studies. *Philosophical Transactions of the Royal Society of London B*, **356**, 351–367.

—— 2007. Marine diversity through the Phanerozoic: problems and prospects. *Journal of the Geological Society, London*, **164**, 731–745.

—— and MCGOWAN, A. J. 2005. Cyclicity in the fossil record mirrors rock outcrop area. *Biology Letters*, **1**, 443–445.

—— ——2007. The shape of the Phanerozoic marine palaeodiversity curve: how much can be predicted from the sedimentary rock record of Western Europe? *Palaeontology*, **50**, 765–774.

—— GALE, A. S. and MONKS, N. E. A. 2001. Sea-level change and rock-record bias in the Cretaceous: a problem for extinction and biodiversity studies. *Paleobiology*, **27**, 241–253.

SPATH, L. F. 1936. The ammonites of the Green Ammonite Beds of Dorset. *Quarterly Journal of the Geological Society*, **92**, 438–455.

TWITCHETT, R. J., WIGNALL, P. B. and BENTON, M. J. 2000. Discussion on Lazarus taxa and fossil abundance at times of biotic crisis. *Journal of the Geological Society, London*, **157**, 511–512.

UHEN, M. D. and PYENSON, N. D. 2007. Diversity estimates, biases, and historiographic effects: resolving cetacean diversity in the Tertiary. *Palaeontologia Electronica*, **10**, 11A–22.

WALL, P. D., IVANY, L. C. and WILKINSON, B. H. 2009. Revisiting Raup: exploring the influence of outcrop area on diversity in light of modern sample-standardization techniques. *Paleobiology*, **35**, 146–167.

WANG, S. C. and DODSON, P. 2006. Estimating the diversity of dinosaurs. *Proceedings of the National Academy of Sciences*, **103**, 13601–13605.

WHALLEY, P. E. S. 1985. The systematics and palaeogeography of the Lower Jurassic insects of Dorset, England. *Bulletin of the British Museum (Natural History), Geology*, **39**, 107–180.

WIGNALL, P. B. and BENTON, M. J. 1999. Lazarus taxa and fossil abundance at times of biotic crisis. *Journal of the Geological Society, London*, **156**, 453–456.

ZEUNER, F. E. 1962. Fossil insects from the Lower Lias of Charmouth, Dorset. *Bulletin of the British Museum (Natural History), Geology*, **7**/5, 155–171.