Biodiversity’s Big Wet Secret: The Global Distribution of Marine Biological Records Reveals Chronic Under-Exploration of the Deep Pelagic Ocean

Thomas J. Webb1*, Edward Vanden Berghe2, Ron O’Dor3

1 Department of Animal and Plant Sciences, University of Sheffield, Sheffield, United Kingdom, 2 Ocean Biogeographic Information System, Institute of Marine and Coastal Sciences, Rutgers University, New Brunswick, New Jersey, United States of America, 3 Census of Marine Life, Consortium for Ocean Leadership, Washington, D. C., United States of America

Abstract

Background: Understanding the distribution of marine biodiversity is a crucial first step towards the effective and sustainable management of marine ecosystems. Recent efforts to collate location records from marine surveys enable us to assemble a global picture of recorded marine biodiversity. They also effectively highlight gaps in our knowledge of particular marine regions. In particular, the deep pelagic ocean – the largest biome on Earth – is chronically under-represented in global databases of marine biodiversity.

Methodology/Principal Findings: We use data from the Ocean Biogeographic Information System to plot the position in the water column of ca 7 million records of marine species occurrences. Records from relatively shallow waters dominate this global picture of recorded marine biodiversity. In addition, standardising the number of records from regions of the ocean differing in depth reveals that regardless of ocean depth, most records come either from surface waters or the sea bed. Midwater biodiversity is drastically under-represented.

Conclusions/Significance: The deep pelagic ocean is the largest habitat by volume on Earth, yet it remains biodiversity’s big wet secret, as it is hugely under-represented in global databases of marine biological records. Given both its value in the provision of a range of ecosystem services, and its vulnerability to threats including overfishing and climate change, there is a pressing need to increase our knowledge of Earth’s largest ecosystem.

Introduction

The tragedy of studying biodiversity during an extinction crisis is that we are losing our subject matter faster than we are able to describe it [1]. This is especially true in the marine environment, where the need to value and conserve taxa and habitats that we know little about has been termed a paradox of marine conservation [2]. Recent efforts by international networks such as the Marine Biodiversity and Ecosystem Functioning EU Network of Excellence (www.marbef.org) and the Census of Marine Life (www.comml.org) have substantially advanced our knowledge of the marine diversity of specific regions [3,4] and habitats [5], in large part by harnessing the power of integrated databases [6]. As well as highlighting what we know about marine biodiversity, however, such databases also allow us to quantify what we do not know. For instance, global synthetic analyses have revealed that even for the best known marine taxa, regional inventories remain worryingly incomplete [7]. Spatial biases are also apparent. In particular, the deep pelagic ocean is revealed as biodiversity’s big wet secret.

The marine pelagic environment is the open oceans and seas, away from the coasts and above the sea bed; and the deep pelagic ocean is typically defined as that part of the water column deeper than 200m. It constitutes a vast biovolume of space in which organisms can exist – by far the largest on Earth at over a billion km² [8-11]. We know that this vast realm and the organisms living in it provide globally important ecosystem services [11], including the support of fisheries, the provision of a range of natural products of potential use in medicine and other applications, as well as the regulation of climate and ocean chemistry through the capture and storage of atmospheric carbon and the production of marine carbonate. But, the limits of our knowledge of this system are continually exposed by the regular discovery of new clades of often large, active and conspicuous organisms [12] whenever surveys are undertaken. Even a charismatic, widely distributed and very large species, the megamouth shark Megachasma pelagios, was not discovered until 1976, and has since been recorded so rarely that each individual specimen has become well known [13].

Although it is generally recognised that our knowledge of the deep pelagic ocean is inadequate [9,11], it is useful to understand
the extent to which this is unusual compared with other regions of the marine environment. Here, we address this question using the most extensive compilation in existence of the spatial distribution of marine taxa. We estimate the location in the water column of c.7 million georeferenced records of marine organisms recorded in the Ocean Biogeographic Information System to provide a graphical summary of the three-dimensional distribution of recorded global marine biodiversity.

**Methods**

**The Ocean Biogeographic Information System**

The objective of OBIS, the Ocean Biogeographic Information System, is to mobilize, integrate and quality control raw biogeography data from many different sources, and to make the resulting data compilation freely and openly available through its international portal (http://www.iobis.org) and other suitable channels [14]. OBIS was created as the data integration component of the Census of Marine Life [15,16], but also holds data originating from its 14 Regional OBIS Nodes, from Thematic Nodes such as FishBase, ICoMM or OBIS SEAMAP, and from independent providers. At present, OBIS makes available through its portal data from well over 700 individual data sets, and has more than 22 million records (a species at a location). It is the largest primary provider of marine biogeographical information, and one of the main providers of data to the Global Biodiversity Information Facility (GBIF). In June 2009, OBIS was adopted by the Intergovernmental Oceanographic Commission as one of its activities under its International Oceanographic Data and Information Exchange programme. Although there are known issues in OBIS with regards to taxonomic completeness, geographical biases, and biogeographic accuracy [7,17], it remains the most complete and comprehensive data repository in existence on the distribution of marine taxa.

**Data extraction and processing**

We extracted all records from OBIS for which sample depth (m) was recorded as well as the latitude and longitude of the sampling event. We used the sample latitude and longitude to determine the depth of the sea floor (bottom depth, m) at that location using ETOPO1 [18]. ETOPO1 is a 1 arc-minute global relief model of Earth’s surface including ocean bathymetry, built from numerous global and regional data sets, and available from http://www.ngdc.noaa.gov/mgg/global/global.html. Together, sample depth and bottom depth describe the position in the water column of each record.

Prior to analysis, we removed 14 sample depths and 1 bottom depth that were negative (most likely intertidal records). Sample depth occasionally exceeded bottom depth (for 7% records), generally by $\leq 20$ m but occasionally by $>100$ m. We make the single reasonable assumption that in all such cases, samples were taken from the sea bed. The discrepancy then either arises from an inaccurate measurement of the depth at which the sample was taken, or due to the geographically less precise estimates of bottom depth (averaged over a grid with a resolution of 1 arc-minute, i.e. approximately 1 nautical mile at the equator) compared to the point measurements of sample depth. Because we cannot distinguish these two possibilities, in all cases we assume the sample depth is correct, and set the bottom depth equal to the sample depth. This has no qualitative bearing on our conclusions. Summing the number of records for each unique combination of bottom depth and sample depth results in a working dataset of 6907676 individual records sampled from 172012 unique locations in three-dimensional (latitude, longitude, sample depth) space. Records were drawn from depths ranging from 0—10670 m, encompassing most of depth range of the global oceans.

**Data Analysis**

We first plot the total number of OBIS records obtained from each distinct bottom depth to provide a global picture of the distribution of recorded marine biodiversity in regions of the ocean of differing depths. Of course, different bottom depths are not equally distributed in the marine environment, and so we next divide the global oceans into five well-recognised zones, based on bottom depth (Table 1). We then used a global sea floor topography [19] to determine the proportional area of the oceans occurring in each zone. For example, the continental shelf ($<200$ m) accounts for around 8.7% of the area of the oceans, whereas the abyssal plain covers almost 50% of the ocean area (Table 1), and is, by surface area, the largest habitat on Earth [9]. Plotting the proportion of OBIS records originating from each zone against the proportional area of that zone shows which areas of the ocean are proportionately under- or over-represented within OBIS.

We next incorporate information on the position within the water column from which each record was sampled (sample depth) to create a three-dimensional picture of recorded marine biodiversity. To add finer resolution to this picture, we further stratified these broad depth zones described in Table 1. Surface waters (0—200 m) were divided into 50 m strata, waters between 200 and 1000 m deep into 100 m strata, waters from 1000 to

---

**Table 1.** Depth zones of the ocean, with the percentage of the global oceans occurring within each zone (estimated from data described in Ref. 17).

| Depth       | Zone                          | Percentage of Ocean Area | Cumulative Percentage | Depth resolution |
|-------------|-------------------------------|--------------------------|-----------------------|-----------------|
| 0—200 m     | A. Continental Shelf          | 8.7                      | 8.7                   | 50 m*           |
| 200—1000 m  | B. Continental Slope / Mesopelagic | 5.8                   | 14.5                  | 100 m          |
| 1000—4000 m | C. Continental Slope / Bathypelagic | 36.3                  | 50.8                  | 200 m          |
| 4000—6000 m | D. Abyssal Plain              | 48.6                     | 99.4                  | 200 m          |
| >6000 m     | E. Hadal Zone                 | 0.6                      | 100                   | 1000 m         |

*Surface waters to 200 m deep were subdivided into 50 m strata for global analyses, but into 10 m strata for analyses of the Continental Shelf and Mesopelagic Continental Slope. The letters (A-E) used for each depth zone match those used in Figures 2 and 3. Depth resolution refers to the subdivision of the water column to each depth zone used in subsequent analyses.

doi:10.1371/journal.pone.0010223.t001
The majority of our knowledge of marine biodiversity is derived from samples drawn from shallow seas: the number of records in OBIS declines precipitously with increasing bottom depth (Figure 1A). Areas of the ocean with a relatively shallow bottom (<200m) typically have thousands of associated records, whereas the deep oceans (>6000m) generally have <10 records. The lowess smooth in Figure 1A indicates that the decline in record numbers is steepest in the range 0—1000m, and again around 5000—6000m. In part this may be related to different depths of ocean having different areas. For instance, the low numbers of records in the hadal zone may be due to the fact that this zone constitutes a very small proportion of the total area of the ocean (Table 1). By plotting the proportion of all records occurring within each of the five ocean regions described in Table 1 against the proportion of ocean area that region encompasses (Figure 1B), the simultaneous effects of proportional area and number of records can be untangled. Thus, although the hadal zone has about the number of records expected, given its tiny area, other zones are either highly under- or over-represented. For instance, >50% of the OBIS records come from the continental shelf, which constitutes <10% of the ocean, whereas <10% of records are from the Abyssal Plain (4000—6000m), which constitutes c. 50% of the ocean area (Figure 1B).

The global distribution through the water column of recorded marine biodiversity is shown in Figure 2. Even on the logarithmic scale of number of records, the dominance of shallower waters within the OBIS database is clear. It is also clear, both throughout the oceans and over the continental shelf and slope, that most records have come from a narrow band of water at the ocean surface, or from the sea bed. The blue area throughout large parts of the deep pelagic ocean in particular reflects the paucity of records from this habitat. When we consider that each cell of 200m depth over the abyssal plain represents a volume of c. 3.5 million km³, the immense volume of water from which biological records are essentially absent becomes starkly apparent.

Figure 1. The depth distribution of OBIS records of global marine biodiversity. A. Number of OBIS records against ocean depth (grey symbols); the general trend is illustrated with a lowess smooth (solid line). Horizontal dashed lines indicate the divisions into regions A-E defined by depth at 200, 1000, 4000 and 6000m (see Table 1). B. The proportion of all OBIS records occurring in the different depth zones identified in Table 1, against the proportion of the global ocean that occurs at those depths. The 1:1 line identifies those areas of the ocean with proportionately more (points above the line) or fewer (points below the line) records than expected given their area. This gives a conservative view of under- and over-representation based on the volume of each habitat.
doi:10.1371/journal.pone.0010223.g001
The general lack of pelagic records is shown more clearly by considering the proportion of records at each position in the water column separately for each bottom depth (see Figure S1). More generally, midwater diversity is under-represented at all depths: in no zone of the ocean do midwater records make up more than 50% of all records (Figure 3). However, this pattern is especially marked in the deep sea: there is a significant negative correlation between bottom depth and the proportion of midwater records (Spearman’s rank correlation, $r_s = -0.66$, d.f. = 54, $P < 0.0001$).

**Discussion**

Our results clearly show that the deep oceans are vastly under-represented in OBIS, the world’s largest provider of information on the distribution of marine species. More generally, midwater diversity is under-represented at all depths: in no zone of the ocean do midwater records make up more than 50% of all records (Figure 3). However, this pattern is especially marked in the deep sea: there is a significant negative correlation between bottom depth and the proportion of midwater records (Spearman’s rank correlation, $r_s = -0.66$, d.f. = 54, $P < 0.0001$).

The general lack of pelagic records is shown more clearly by considering the proportion of records at each position in the water column separately for each bottom depth (see Figure S1). More generally, midwater diversity is under-represented at all depths: in no zone of the ocean do midwater records make up more than 50% of all records (Figure 3). However, this pattern is especially marked in the deep sea: there is a significant negative correlation between bottom depth and the proportion of midwater records (Spearman’s rank correlation, $r_s = -0.66$, d.f. = 54, $P < 0.0001$).

There are two possible explanations for this: either the deep pelagic ocean is especially low in biomass, or it has been especially under-sampled (or some combination of the two). Historically, the first of these possibilities has been espoused. For instance, Charles Wyville Thomson, leader of the Challenger Expedition in the 1870s which effectively launched the discipline of deep sea biology [11], believed that ‘the fauna of deep water is confined primarily to two belts, one at and near the surface and the other on and near the bottom; leaving an intermediate zone [i.e., the deep pelagic] in which larger animals… are nearly or entirely absent’ (quoted in Ref. 9). More recent evidence suggests, however, that it is under sampling and net avoidance rather than a lack of organisms that generate the patterns we have observed.

However, new technologies have dramatically altered perceptions of the deep pelagic ecosystem [10,21], suggesting that with past techniques, even high sampling effort may not have resulted in correspondingly high numbers of biological specimens being collected. The ability to view animals *in situ* means that the diversity of organisms not captured by traditional sampling methods, such as the gelatinous fauna that constitutes up to a quarter of pelagic biomass [10] is now better understood. Importantly, their abundance is now known to be much higher than most deep-sea biologists expected [10]. For instance, the recently discovered new clade of large, active deep sea annelids
(including holopelagic species) occur at high biomass [12]. Thus, the deep pelagic appears to conform to dictum that the more you survey, the more you find, as witnessed recently in other marine habitats including fish in the hadal zone [22] and microbes in surface waters [23,24]. Such findings have led to estimates of a million undescribed species in the deep pelagic [10] and the proposal that ‘within this vast midwater habitat are the planet’s largest animal communities… These animals probably outnumber all others on Earth’ (Ref. 10:848). Clearly, there is much work still to be done before we can draw conclusions regarding the depth distribution of actual marine biodiversity from databases of recorded marine biodiversity.

Increasing our understanding of these communities is important for a number of reasons. First are the ecosystem services they provide, for instance supporting global fisheries, climate regulation, and bioprospecting [11]. In addition, they have considerable potential as a model system for testing biogeographic hypotheses, such as large-scale gradients in diversity. The deep pelagic environment is spatially homogeneous and has been very stable over time, with little in the way of seasonal and latitudinal variability [10], and yet latitudinal gradients appear to exist in the diversity of at least some deep pelagic taxa [8]. Might this provide a means to tease apart the confounding effects of the environment, geometric constraints, and species tolerances in explaining biogeographic patterns [25,26]? More generally, it may prove easier to unravel the multiple drivers of change in marine ecosystems, including historical human influences and future climate change, by studying those habitats that have been least affected to date – the mid-ocean, mid-water environment – before transferring this understanding back into more heavily disturbed coastal and benthic systems [27].

Finally, even if pelagic ecosystems remain less impacted than coastal regions, there is increasing concern that human activities including fishing, pollution and climate change have already had substantial effects, and that these pressures are only likely to increase in future [11,28-30]. Although some conservation measures, in particular the establishment of pelagic marine protected areas, may be possible in the absence of detailed biological information [28], clearly an increased understanding of the temporal and spatial dynamics of pelagic organisms will improve their effectiveness. We hope that exposing biodiversity’s big wet secret will stimulate further exploration of Earth’s biggest ecosystem.

**Supporting Information**

**Appendix S1**  R code used for data processing and production of all figures.
Figure S1 Global distribution of recorded marine biodiversity expressed as the proportion of OBIS records occurring at each position in the water column over each bottom depth. Only those cells (unique combinations of sample and bottom depth, following the scheme in Table 1) contributing more than 1% of records from a given bottom depth are coloured. To provide an alternative scale to Figure 2, where the size of each cell represents the volume of water it contains, here we transform both the vertical depth axis (d) and the horizontal area axis (a) \( \text{d}^2 = \text{d}^{2/3}, \ a^2 = \text{a} \). This better shows the relatively high number of records in surface waters, as well as patterns in the smaller depth zones (e.g., continental shelf, hadal zone).

References

1. Holt RD (2010) 2020 Visions: Ecology. Nature 463: 32.
2. Webb TJ (2009) Biodiversity research sets sail: showcasing the diversity of marine life. Biol Lett 5: 145–147.
3. Somerfield PJ, Arvanitidis C, Vanden Berghe E, van Averbeke P, Hummel H, et al. (2009) MacBef, databases, and the legacy of John Gray. Mar Ecol Prog Ser 382: 221–224.
4. Vanden Berghe E, Claus S, Appelhans W, Faulwetter S, Arvanitidis C, et al. (2009) MacRen, integrated database on benthic invertebrates of European continental shelves: a tool for large-scale analysis across Europe. Mar Ecol Prog Ser 382: 225–238.
5. Clark MR, Rowden AA, Schlacher R, Williams A, Consalvey M, et al. (2010) The ecology of seamounts: structure, function, and human impacts. Annu Rev Marine Sci 2: 253–278.
6. Costello MJ, Vanden Berghe E. (2006) ‘Ocean biodiversity informatics’: a new era in marine biology research and management. Mar Ecol Prog Ser 316: 203–214.
7. Mora C, Tittensor DP, Myers RA (2008) The completeness of taxonomic inventories for describing the global diversity and distribution of marine fishes. Proc R Soc B 275: 149–155.
8. Angel MV (1993) Biodiversity of the pelagic ocean. Cons Biol 7: 760–772.
9. Kunzig R (2003) Deep-sea biology: living with the endless frontier. Science 302: 991.
10. Robinson BH (2004) Deep pelagic biology. J Exp Mar Biol Ecol 300: 253–272.
11. Robinson BH (2009) Conservation of deep pelagic biodiversity. Cons Biol 23: 847–858.
12. Osborn KJ, Haddock SHD, Pleijel F, Madin LP, Rous SW (2009) Deep-sea, swimming worms with luminescent ‘bombs’. Science 325: 964.
13. Berra TM (1997) Some 20th Century fish discoveries. Environ Biol Fish 50: 1–12.
14. Vanden Berghe E, Graas J, Stocks K, Halpin P, Lang da Silva Veira F (2010) Integrating Biological Data Into Ocean Observing Systems: The Future Role of OBIS. In: Hall J, Harrison DE, Stammer D, eds. Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2), Venice, Italy, 20–25 September 2009, ESO Publication WPP-306. Available: http://www.oceanobs09.net/blog/?p=271. Accessed 2010 Feb 1.
15. Forrester M (2000) Planning for OBIS: examining relationships with existing national and international biodiversity information systems. Oceanography 13: 31–38.
16. Grasdo JF (2000) The Ocean Biogeographic Information System (OBIS): an online, worldwide atlas for accessing, modelling and mapping marine biological data in a multidimensional geographic context. Oceanography 13: 5–7.
17. Robertson DR (2008) Global biogeographical databases on marine fishes: caveat emptor. Divers Distrib 14: 891–892.
18. Amante C, Eakins BW (2009) ETOP01 1 arc-minute global relief model: procedures, data sources and analyses. NOAA Technical Memorandum NESDIS NGDC-24.
19. Smith WHF, Sandwell DT (1997) Global sea floor topography from satellite altimetry and ship depth soundings. Science 277: 1956–1962.
20. R Development Core Team (2009) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.
21. Haddock SHD (2004) A golden age of gelata: past and future research on planktonic ctenophores and cnidarians. Hydrobiologia 530/531: 549–556.
22. Jamieson AJ, Fujii T, Solan M, Matsumoto AR, Bagley PAJ, et al. (2008) Liparid and macrourid fishes of the hadal zone: in situ observations of activity and feeding behaviour. Proc R Soc B 276: 1037–1041.
23. Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, et al. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. Science 304: 66–74.
24. Pommer T, Canback B, Riemann I, Boström KH, Simu K, et al. (2007) Global patterns of diversity and community structure in marine bacterioplankton. Mol Ecol 16: 867–880.
25. Storch D, Davies RG, Zajicke R, Orme CDL, Olson V, et al. (2006) Energy, range dynamics and global species richness patterns: reconciling mid-domain effects and environmental determinants of avian diversity. Ecol Lett 9: 1308–1320.
26. Siddig AL, Storch D, Keil P (2009) Rapoport’s rule, species tolerances, and the latitudinal diversity gradient: geographic considerations. Ecology 90: 3575–3586.
27. O’Dor R, Acosta J, Bergstad OA, Brainard R, Brattey J, et al. (2010) Bringing life to ocean observation. In: Hall J, Harrison DE, Stammer D, eds. Proceedings of OceanObs’09: Sustained Ocean Observations and Information for Society (Vol. 2), Venice, Italy, 21–25 September 2009, ESA Publication WPP-306. Available: http://www.oceanobs09.net/blog/?p=310. Accessed 2010 Feb 1.
28. Game ET, Grantham HS, Hobday A, Pressey RL, Lombard AT, et al. (2009) Pelagic protected areas: the missing dimension in ocean conservation. Trends Ecol Evol 24: 360–369.
29. Fabry VJ, Seibel BA, Feely RA, Orr JC (2008) Impacts of ocean acidification on marine fauna and ecosystem processes. ICES J Mar Sci 65: 414–432.
30. Riebesell U, Koertzinger A, Oschlies A (2009) Sensitivities of marine carbon fluxes to ocean change. Proc Natl Acad Sci USA 106: 20662–20669.

Acknowledgments

This publication is considered a contribution to the Census of Marine Life. The authors also are grateful to the many data custodians who have been willing to share their data with OBIS, thus making possible this type of large-scale analysis.

Author Contributions

Conceived and designed the experiments: TW EVB RO. Performed the experiments: TW EVB. Analyzed the data: TW EVB. Contributed reagents/materials/analysis tools: TW EVB RO. Wrote the paper: TW EVB RO.

Found at: doi:10.1371/journal.pone.0010223.s001 (0.02 MB PDF)

Biodiversity’s Big Wet Secret