doi:10.1038/nature12529

Quality control, biases and confidence in fish data

The extent and quantity of data used for this study and in the RLS database has only been possible through volunteer efforts. Rigorous quality control measures have been applied to ensure consistency and quality of data, and data have not been contributed by every diver willing to participate – interested participants were vetted on experience and capability. Despite being built on volunteer efforts, the database includes data collected by a combination of experienced scientists and trained recreational RLS divers. For example, over 60% of the data used in this global analysis came from divers with at least tertiary level scientific training, and 34% came from divers with PhD qualifications.

For all RLS surveys, data are transcribed from underwater datasheets into custom data entry forms in Microsoft Excel, usually the same day as dives are undertaken. Excel data entry templates contain lookups from region-specific species lists, have a number of automatic data checks and are in a consistent format for uploading to the RLS database. Further data checks are made upon upload to the database, including for data structure and consistency in meta-data among divers, as well as checks designed to detect species which have not previously been recorded in that particular region (which prompt querying particular data points; and checking taxonomic and distributional data before addition of new species).

RLS field campaigns, such as those used for data collection for this study, typically involve a small group of divers undertaking survey dives over a period of three days to two weeks (or occasionally longer), under the direction and supervision of a scientist or experienced survey diver. At the conclusion of field campaigns, one of the RLS organisers or scientists leading the field trip collates data from participants and additionally undertakes many of these checks manually. This includes close scrutiny of species lists, abundances and site details. Evidence in the form of images is typically requested for records of species which are not seen by the experienced surveyor on the trip, with such evidence essential for divers with less experience in that particular region. Uncertain records or records of new species for regions for which definitive evidence does not exist are reduced to the highest taxonomic resolution for which there is certainty (usually genus).

For divers without prior formal scientific training, the vast majority who contributed data to this study were trained by the same two authors (RDSS and GJE). Previous comparison of data collected by those RLS divers without formal scientific training with those of experienced scientists has shown that the variation between divers is negligible in comparison to the variation among and within sites[43]. Thus, minor differences in skill and experience among divers are unlikely to affect the broad-scale relationships identified in this study. Furthermore, authors GJE and RDSS participated in surveys over many ecoregions (60 of 74 between them), providing a substantial element of consistency in diver participation at the global scale.

The dataset used in this analysis included 4,357 transects at 1,844 sites in 74 marine ecoregions and 11 realms[44]. Time-series information at sites from repeat surveys as part of ongoing RLS monitoring in Australia and Spain was excluded from this analysis, which focuses on spatial patterns. Within-site transect replication varied between sites, with a global average of 2.4 (±0.04 SE) transects per site (therefore an average coverage of 1,200m^2 per site). While variable, and occasionally low, within-site replication may affect the accuracy in characterisation of the fish fauna at individual sites, this variation is unlikely to influence patterns detected at the global scale. Site was the level of replication in our global analysis, and analysis of 1,844 sites with such a broad latitudinal distribution (133 degrees) represents exceptional replication for a study of this nature.

As for all methods for quantifying fish communities, visual census methods are characterised by large amounts of variation and a number of biases (e.g. see[45]). We make the assumption that none of these biases are directional in such a way that will influence the values of diversity metrics over the global scale - for example, that species-specific behavioural traits which affect estimates of their density will be consistent throughout their range.

Allocation of functional traits and potential bias

Allocation of trait values was undertaken for the species list as a whole, without any distinction or division related to species' geographic distributions, or any other reason for spatial bias in terms of human allocation. Bias could still arise in the trait database through differences in the accuracy of information available through FishBase, which provided information used in combination with other published work for the maximum size, trophic breath, and trophic group traits. We recognise that FishBase values are of varying reliability, and in each case FishBase values were checked against the expectations of the authors after population of the trait values to the whole list. For example, maximum size values taken from FishBase were cross-checked, and if not consistent with authors' expectations, a web search was undertaken to find any alternative published maximum length values that appeared more accurate according to the authors' knowledge of the species. Likewise, trophic breadth and trophic group were populated through a combination of dietary studies reported in FishBase, supplemented by an extensive literature search by one of the authors (GAS) as part of another study. Trophic group was allocated based on a combination of FishBase information and this dietary dataset.

The remaining traits are unlikely to vary spatially for each species. For example, there is little error in classifying a species for the gregariousness or water column position traits. These can be obtained from *in situ* observation and tend to be highly consistent for species across their entire range.

Appropriateness of diversity indices

Species density, expressed as the number of species observed per 500 m^2 transect area, was considered an appropriate measure of species richness in our study for comparison of reef sites across different parts of the world. While surveys will not cover the species present in a broader region in equal proportions (due to the varying nature of species-area curves over a global scale), we employed fixed-coverage sampling[46] to calculate rarefied estimates of richness at 90, 95 and 99% coverage to assess differences in coverage of species present at the site scale. This process identified that 98.7% of sites had greater than 90% coverage in terms of the site-level number of species and our species density measure was closely related to estimates at all three levels of coverage tested (Pearson's r = 0.87, 0.86 and 0.89, respectively). In other words, continued sampling at the vast majority of sites would have a relatively low probability of including additional species.

With respect to interpretation, our measure of richness represents a realistic and ecologically-meaningful estimate of the number of species living together at the site scale, more truly reflecting α diversity at this scale. Traditional species richness measures based on aggregated occurrence records over a broader scale include many species which may not actually live together on a given patch of reef (and thus potentially confound α and β diversity components), and are also characterised by vastly greater variability in search/survey effort between regions. Despite these differences, our patterns in species density closely match previously reported species richness patterns for marine fishes[47,48], and thus both methods appear comparable over the global scale.

We additionally recalculated our functional diversity and evenness metrics by pooling data across two transects (1000 m^2), randomly drawn from sites where more than two transects were surveyed, to explore whether inclusion of more rare species through calculation of metrics over a greater area within each site (compared to averaging values per 500 m^2) greatly affected site level estimates. New estimates were highly correlated to those we report based on mean values per 500 m^2 (Pearson's r = 0.90 and 0.88 for functional diversity and evenness, respectively), and thus we report findings based on means per 500 m^2 and using the full dataset.

REFERENCES

43 Edgar, G. J. & Stuart-Smith, R. D. Ecological effects of marine protected areas on rocky reef communities: a continental-scale analysis. *Mar. Ecol. Prog. Ser.* **388**, 51-62 (2009).

44 Spalding, M. D. *et al.* Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *Bioscience* **57**, 573-583 (2007).

45 Edgar, G. J., Barrett, N. S. & Morton, A. J. Biases associated with the use of underwater visual census techniques to quantify the density and size-structure of fish populations. *J. Exp. Mar. Biol. Ecol.* **308**, 269-290 (2004).

46 Chao, A. & Jost, L. Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology* **93**, 2533-2547 (2012).

47 Allen, G. R. Conservation hotspots of biodiversity and endemism for Indo-Pacific coral reef fishes. *Aquatic Conserv: Mar. Freshw. Ecosyst.* **18**, 541-556 (2008).

48 Tittensor, D. P. *et al.* Global patterns and predictors of marine biodiversity across taxa. *Nature* **466**, 1098-1101 (2010).