

## FEATURE PAPER

**Setting evolutionary-based conservation priorities for a phylogenetically data-poor taxonomic group (Scleractinia)**

D. J. Curnick<sup>1,2</sup>, C. E. I. Head<sup>2,3</sup>, D. Huang<sup>4</sup>, M. J. C. Crabbe<sup>3,5</sup>, M. Gollock<sup>2</sup>, B. W. Hoeksema<sup>6</sup>, K. G. Johnson<sup>7</sup>, R. Jones<sup>2</sup>, H. J. Koldewey<sup>2</sup>, D. O. Obura<sup>8</sup>, B. R. Rosen<sup>7</sup>, D. J. Smith<sup>9</sup>, M. L. Taylor<sup>3</sup>, J. R. Turner<sup>10</sup>, S. Wren<sup>2,11</sup> & D. W. Redding<sup>1</sup>

1 Centre for Biodiversity and Environment Research, Department of Genetics, Evolution and Environment, University College London, London, UK

2 Zoological Society of London, London, UK

3 Department of Zoology, University of Oxford, Oxford, UK

4 Department of Biological Sciences, National University of Singapore, Singapore

5 Institute of Biomedical and Environmental Science and Technology, Faculty of Creative Arts, Technologies and Science, University of Bedfordshire, Luton, UK

6 Department of Marine Zoology, Naturalis Biodiversity Center, Leiden, The Netherlands

7 Department of Earth Sciences, Natural History Museum, London, UK

8 Coastal Oceans Research and Development in the Indian Ocean (CORDIO) East Africa, Mombasa, Kenya

9 Coral Reef Research Unit, University of Essex, Colchester, UK

10 School of Ocean Sciences, Bangor University, Anglesey, UK

11 Department of Zoology, University of Otago, Dunedin, New Zealand

**Keywords**

coral; data-poor; evolutionary distinctiveness; ED score; prioritization; phylogenetic diversity; Scleractinia.

**Correspondence**

David J. Curnick, Centre for Biodiversity and Environment Research, University College London, Gower Street, London WC1E 6BT, UK.

Email: d.curnick@ucl.ac.uk;

david.curnick@zsl.org

Catherine E. I. Head, Department of Zoology, University of Oxford, Tinbergen Building, South Parks Road, Oxford OX1 3PS, UK.

Email: catherine.head@zoo.ox.ac.uk

Editor: Res Altwegg

Associate Editor: Jonathan Davies

Received 16 June 2013; accepted 10 November 2014

doi:10.1111/acv.12185

**Introduction**

Current species extinction rates across all taxa are overwhelming background rates and could increase further (Pimm *et al.*, 2006; Barnosky *et al.*, 2011; De Vos *et al.*, 2014; WWF 2014). With high numbers of species under threat of extinction and a global shortfall in funding (Waldron *et al.*, 2013), there is a pressing need to establish

**Abstract**

Given the current extinction crisis coupled with the shortfall in funding, there is a pressing need to establish species conservation priorities. The prioritization of phylogenetic diversity and evolutionary distinctiveness is one approach; however, taking such an approach requires more phylogenetic data than are currently available for most taxa. Here, we investigate the effects of increased phylogenetic knowledge on the accuracy of evolutionary distinctiveness (ED) scores over time using scleractinian corals as a case study. ED scores were calculated from four molecular-based phylogenies from 2008 to 2013, each one representing a chronological step of increased phylogenetic knowledge for scleractinian corals, finally resulting in a full species-level phylogeny which is used here as the reference dataset. As expected, the most complete and up-to-date phylogenies performed well at predicting scores taken from a recent, full-coverage species-level phylogeny of scleractinian corals. Surprisingly, however, older phylogenies and scores derived from expert opinion also performed well. More unexpectedly, the expert opinion-led scores, when used as a basis for imputing scores for missing species, achieved a close second in terms of prediction accuracy compared with the most recent and largest tree, which had nearly 10 times more taxonomic coverage. We recommend, once tested further, that ED score imputation be considered for assessing the conservation priorities for other poorly studied groups.

conservation priorities among threatened species (e.g. Bibby *et al.*, 1992; Mittermeier *et al.*, 1998; Olson & Dinerstein, 1998; Myers *et al.*, 2000; O'Connor, Marvier & Kareiva, 2003; Possingham & Wilson, 2005; Brooks *et al.*, 2006; Margules & Sarkar, 2006; Isaac *et al.*, 2007).

A large number of conservation priority-setting approaches have been documented in the literature (e.g. Brooks *et al.*, 2006; Wilson *et al.*, 2006; Pressey *et al.*, 2007;

Treblico *et al.*, 2011; Selig *et al.*, 2014). One species-based method considers the evolutionary relationships within a clade of target species (Vane-Wright, Humphries & Williams, 1991; Faith, 1992, 1996, 2007) alongside threat status to set conservation priorities (Witting & Loeschcke, 1995; Redding & Mooers, 2006; Isaac *et al.*, 2007, 2012; Safi *et al.*, 2013). The goal of prioritizing evolutionarily distinct species is to maximize the number of evolutionary trajectories conserved (Vane-Wright *et al.*, 1991; Crozier, 1992). This precautionary, maximization approach is necessary as future evolutionary radiations are difficult to predict (Cooper & Fortey, 1998), which makes picking evolutionary 'winners' extremely difficult. Conservation actions based on measures of species-specific evolutionary isolation have been shown to conserve a greater proportion of the overall phylogenetic diversity of the target clade (Redding *et al.*, 2008) and can preserve some unusual life-history and morphological traits, potentially helping to protect other aspects of biodiversity within a clade (Cadotte, Cardinale & Oakley, 2008; Cadotte *et al.*, 2009; Redding, Dewolff & Mooers, 2010). A conservation programme that has applied this approach is the Evolutionarily Distinct and Globally Endangered (EDGE) of Existence programme (Isaac *et al.*, 2007; <http://www.edgeofexistence.org>). The EDGE ranking system is a composite score based on two components: a measure of evolutionary distinctiveness (ED) and a measure of global endangerment (GE) or threat (Redding & Mooers, 2006; Isaac *et al.*, 2007; Mooers, Faith & Maddison, 2008). To date, this has been applied to mammals, amphibians and birds (Isaac *et al.*, 2012; Safi *et al.*, 2013; Jetz *et al.*, 2014).

While it may be the case that evolutionary-based conservation approaches offer a data-cheap method of assessing several aspects of biodiversity for the target clade concurrently (Redding *et al.*, 2010), taking such an approach has only been attempted with data-rich species groups (e.g. mammals, birds) as it requires more phylogenetic data than are currently available for most taxa. Indeed, little work has been carried out to test how phylogenetic priority setting can be applied to relatively data-poor taxonomic groups, in particular those that have not been subjected to the same intensive regimes of genetic sequencing carried out for large-bodied vertebrates.

The threats to tropical coral reefs have been reported extensively (Sheppard, 2003; Donner *et al.*, 2005; Hoegh-Guldberg *et al.*, 2007; Veron *et al.*, 2009; Dove *et al.*, 2013; Kennedy *et al.*, 2013). In 2006 and 2007, the threat status of the 827 reef-building scleractinian species were assessed by the world's leading coral experts for the International Union for the Conservation of Nature (IUCN) Red List (Carpenter *et al.*, 2008). Of the 688 species not deemed to be 'data deficient', 32.7% were listed in an elevated threat category ('vulnerable', 'endangered', 'critically endangered'). In 2010, the Zoological Society of London set evolutionary-based conservation priorities for scleractinian corals, a group which, at that time, lacked detailed phylogenetic data. The assessment was based on available phylogenies (Kerr, 2005; Fukami *et al.*, 2008) and taxonomy, combined with expert opinion (as outlined in the

succeeding text; hereafter called the 2010 EDGE Coral Expert Workshop). A later iteration of EDGE scores was then published based on a more recent species-level phylogenetic supertree (Huang, 2012). This provides a unique opportunity to validate the expert-led approach, while at the same time investigate how limited phylogenetic knowledge may affect our confidence with which species are chosen as priority species for conservation attention at each stage of assessment within the EDGE framework.

Here, for the first time, we investigate setting evolutionary-based conservation priorities for a previously data-poor species group: scleractinian corals. We examine the role of increasing phylogenetic information on improving the accuracy of ED scores using five empirically derived datasets that represent chronological steps of increasing phylogenetic coverage: (1) 127 species two-gene phylogeny (Fukami *et al.*, 2008); (2) Expert opinion added to the previous 127 species phylogeny; (3) 234 species single-gene phylogeny (Kitahara *et al.*, 2010); (4) 466 species multi-gene phylogeny (Huang & Roy, 2013); (5) 838 species supertree (Huang, 2012). For each of these datasets, we determined the consistency of top 50 scoring species and then measured the correlation of the ED scores to the most recent 'reference' score set, in this case the high coverage supertree (5, above). Finally, we explore an alternative method of increasing the phylogenetic coverage of evolutionary-led conservation priorities using taxonomic imputation methods to estimate the distinctiveness scores of unsampled taxa. We show that setting conservation priorities is possible for groups with only limited sampling and that imputation offers a pragmatic method to estimate scores. Our approach has the potential to open up evolutionary-led conservation methods to the vast majority of species groups for which there is limited genetic sampling.

## Materials and methods

We first surveyed the literature to determine how phylogenetic knowledge on scleractinian corals has developed since 2008 and how this expanding knowledge would have altered the EDGE-based conservation priorities within this group over time. We isolated five points in time where coral phylogenetic knowledge had notably increased and at each stage calculated ED and EDGE scores using these data. The points in time were as follows:

### 2008 – 127 species four-gene phylogeny

Fukami *et al.* (2008) analysed 127 species across 75 genera and 17 families using two mitochondrial DNA genes and two protein-coding nuclear genes in both maximum likelihood and Bayesian analyses, returning relatively strong support and high resolution for many major groups. To conform to the previous iterations of the EDGE process, we dated the root of Scleractinia at 425 million years ago (Stolarski *et al.*, 2011) and used a strict molecular clock (MrBayes 3.2.1; Huelsenbeck & Ronquist, 2001; Ronquist

& Huelsenbeck, 2003) to create a time-based ultrametric tree (10 million generations, burn-in = 25%, fixed topology).

### 2010a – EDGE Coral Expert Workshop

In the absence of a complete phylogeny, the EDGE programme held a participatory workshop (EDGE Coral Reefs Workshop, 4–5 March 2010, Zoological Society of London) to identify 10 priority EDGE species with which to initiate the EDGE Corals conservation programme. The workshop was attended by 22 scientists with a further 10 consulted remotely.

Prior to the workshop, 51 species were shortlisted for consideration through remote consultation with a range of coral experts (Supporting Information Appendix S1) and from ED scores derived from the Fukami *et al.* (2008) phylogeny. This phylogeny, although acknowledged to be incomplete, was used as the key reference during the 2010 EDGE Coral Expert Workshop as being the best available information at the time. A further five species were added during the initial workshop review process, thus totalling 56 species (Supporting Information Appendix S1). Participants were separated into two isolated working groups and asked to quantitatively score each species' ED (arbitrary scoring of 1 to 4, with 4 being the highest ED) based on phylogenetic relationships presented in the Fukami *et al.* (2008) phylogeny and expert opinion against set criteria (Supporting Information Appendix S2). Note that no new phylogenies were reconstructed based on primary data for these 56 species.

### 2010b – 234 species single-gene barcode phylogeny

Published in late 2010 (Kitahara *et al.*, 2010), this phylogeny had greater taxonomic coverage, but only used a single gene, one of the four utilized by Fukami *et al.* (2008): cytochrome c oxidase subunit I (COI), which we note is known to cause potential problems for accurate evolutionary inference (Castresana, 2007). For anthozoans in particular, the rate of COI sequence evolution is extremely low (Shearer *et al.*, 2002; Hellberg, 2006; Huang *et al.*, 2008), reducing the topological resolution of this tree over that of Fukami *et al.* (2008). Of the 21 major clades recognized in the latter, Kitahara *et al.* (2010) recovered 17, although most of the deeper relationships were concordant between them. One of the major clades was displaced by the conflicting placement of *Oulastrea crispata* (clade XI), while the remaining three were unresolved (clades XVII, XIX and XX) owing to low sequence variability of COI (Huang *et al.*, 2009).

Here, we matched species names to those noted within the supertree and in agreement with the taxonomy used by the IUCN (2013) in the red-listing process, resulting in 234 species for which ED scores could be calculated. Again, we used a strict molecular clock (MrBayes 3.2.1) to create a time-based ultrametric tree (20 million generations,

burn-in = 25%, fixed topology) with the root set at 425 million years.

### 2013 – 466 species multi-gene phylogeny

In 2013, a third phylogeny was inferred using an incomplete molecular supermatrix analysis of seven mtDNA markers (Huang & Roy, 2013). This study added three more markers over those used in Fukami *et al.* (2008) and nearly doubled the taxon sampling of Kitahara *et al.* (2010) because of the utilization of genes other than COI. All 21 major clades defined by Fukami *et al.* (2008) were recovered in this phylogeny, which also resolved the deepest three-clade topology ('basal', complex and robust groups; Stolarski *et al.*, 2011) that were not found in the earlier studies.

We note that this product was not optimized by all authors for purposes here but is included as one possible method open to other groups where there is patchy sequencing but no published phylogeny. As for previous iterations, species names from this phylogeny were matched to those within the supertree resulting in 466 Scleractinia species that were in agreement with the IUCN taxonomy (IUCN, 2013).

### 2012 – 838 species supertree – reference tree

The reference tree used here was based on phylogenetic reconstructions published in Huang (2012) and Huang & Roy (2013). The topology of the tree was inferred using a supertree method (Baum, 1992; Ragan, 1992) that incorporated data from 975 reef and non-reef building coral terminals. Reef species constituted 838 of these taxa, including the 827 species evaluated for the IUCN Red List of Threatened Species (Carpenter *et al.*, 2008). The source trees comprised a molecular tree of seven mtDNA genes from 466 species based on the above multi-gene phylogeny (Huang & Roy, 2013), 13 morphological trees and a taxonomic tree. These were coded into a bootstrap-weighted matrix representation and analysed under the maximum parsimony framework to infer a strict consensus supertree. Time-calibrated trees (chronograms) were then generated based on a mtDNA tree calibrated as per Simpson *et al.* (2011) and Stolarski *et al.* (2011). Using BEAST 1.6.2 (Drummond & Rambaut, 2007), five Markov chain Monte Carlo (MCMC) analyses of 30 million generations were carried out with a sampling interval of 1000 and a post-convergence burn-in of 10 million generations. Effective sample size values for all estimated parameters exceeded 200 when the five MCMC runs were combined, although nearly all exceeded 2000. Polytomies on the strict consensus supertree were randomly resolved using PolytomyResolver (Kuhn, Mooers & Thomas, 2011), constraining each node with a normal prior based on the 95% highest posterior density derived from the ultrametricized mtDNA tree (see Collen *et al.*, 2011). Five MCMC runs of 3 million generations were performed and, following rejection of the first 1 million

steps, were subsampled every 10 000 iterations, resulting in 1000 fully-resolved chronograms.

This dataset of 1000 chronograms is the most recent large-scale analysis of the phylogenetic relationships of corals. It contains the largest amount of information in terms of number of sources and relative quantity and also captures some measure of the variability of evolutionary relationships suggested by the data. As a result, we set this as our reference dataset. However, we acknowledge that these evolutionary trees are unlikely to represent the ‘true’ tree as there is still much work to be done to sample and analyse phylogenetic data from scleractinian corals before a phylogenetic consensus is reached (Wiens & Tiu, 2012).

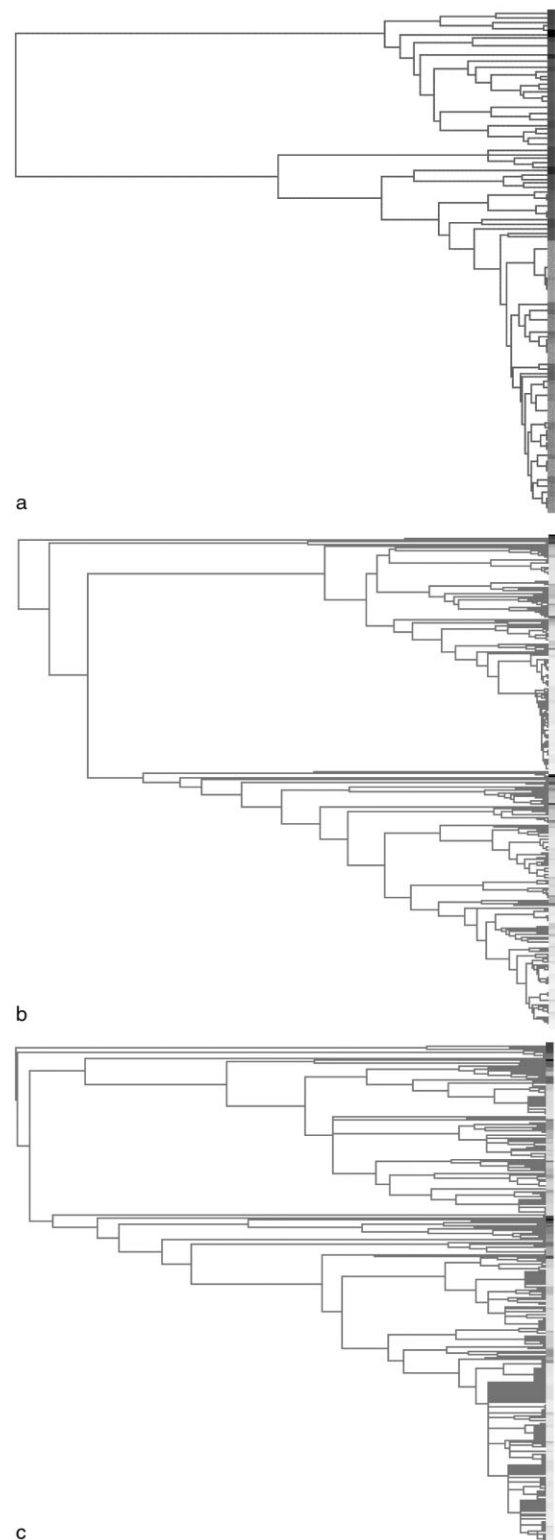
### Null expectation

We created a null expectation to which we compared the older incomplete empirically derived phylogenies by examining how ED-based information on the reference phylogeny degraded as tips were pruned from the complete trees. To create the dataset to run this analysis, we randomly chose a full tree (from the published distribution of 1000) and randomly subsampled this tree to create subtrees that contained 15 to 815 tips incrementally in steps of 50 tips (as a compromise between computational cost and resolution) to give a total of 16 subtrees of increasing size. This process was then repeated 10 000 times, again this number was chosen as a balance between computational cost and examining potential variation between full trees in the distribution and the effect of randomly removing taxa on subtrees. We note this dataset represents a conservative null expectation as we would expect actual subsamples of a phylogeny to capture phylogenetic information and perform better than trees created using different data sources.

### Analysis

For each subtree (i.e. trees in our dataset with fewer than 838 species) taken from either the empirical data or through subsampling the ‘reference’ tree, we calculated the ED metric (Isaac *et al.*, 2007) for all species (Fig. 1). This score divides up all the branches in a target phylogeny among the tips, such that species with the fewest relatives have the largest summed branch length. We then calculated these scores on all versions of the full reference tree to provide a ‘reference’ ED score set (using *ed.calc*, *caper* package in R; Orme *et al.*, 2012). We then calculated the correlation between these subtree scores and the reference set using only complete pairwise observations (Pearson’s rho; *cor* function in R package *base*; R Core Team, 2013).

Then, for all subtrees, we created a full set of 838 ED scores. Scores for missing species, according to taxonomy by Veron (2000), were estimated through statistical imputation (Little & Rubin, 1987; Holland *et al.*, 2006; Swenson, 2013) using scores from related species represented in the subtrees. Specifically, missing representatives of each genus had the



**Figure 1** Molecular phylogenies for Scleractinia corals used in this study to estimate evolutionary distinctiveness scores: Fukami *et al.* 2008 (a), Kitahara *et al.* 2010 (b) and Huang & Roy 2013 (c). Colour strip to the right of each phylogeny indicates evolutionary distinctiveness (ED) score of the tips with darker grey indicating high relative ED score.



mean genus score attributed to all non-sampled members. For a missing species that had no con-generic species, the mean family member score was used instead. If there were no family members represented in the subset tree, the mean ED score of all corals species in the subtree was attributed to the missing species. The development of new molecular techniques has led to new insights on coral evolution and more accurate phylogenetic inferences (e.g. Benzoni *et al.*, 2007; Nunes *et al.*, 2008; Huang *et al.*, 2009; Budd *et al.*, 2010; Gittenberger, Reijen & Hoeksema, 2011; Arrigoni *et al.*, 2012). In fact, our reference tree was built based on data from these recent studies. Therefore, the imputation of scores for missing species based on Veron's (2000) taxonomic classification should be considered a coarse approach and we can ask how our results align given the uncertain taxonomy.

For each subtree, we also created an EDGE species list. GE was measured using the IUCN Red List categories and criteria (IUCN, 2013), widely accepted as the best quantitative and objective framework available for evaluating extinction risk (e.g. Rodrigues *et al.*, 2006; Hoffmann *et al.*, 2008; Mace *et al.*, 2008). Each category was re-coded to give a numeric value: critically endangered = 4, endangered = 3, vulnerable = 2, near threatened = 1, with lower rankings not scored. Threat data for 827 reef building corals were derived from the work of Carpenter *et al.* (2008). For each of the five empirical datasets, we created a list of the top 50 highest EDGE scoring species and then compared membership of the first four iterations with the top 50 EDGE species specified in our reference phylogeny.

## Results

For both pairwise scores and imputed full scores, the ability for subsampled phylogenetic data to predict ED scores from the complete dataset increased in an approximately logarithmic fashion as completeness increased (Fig. 2). For example, half-sized trees (with 419 tips) produced pairwise correlation scores of  $0.773 \pm 0.186$  and imputed full score sets that were  $0.623 \pm 0.368$  (Pearson's rho,  $n = 1000$ ) correlated to the 'reference' set.

The ED scores from the largest, 2013, most recent tree were the most closely correlated (average 0.9,  $SD = 0.01$ ,  $n = 366$ , Spearman's rho; 100 replications) to the reference scores, followed by the 234 species on the 2010b phylogeny (0.7,  $SD = 0.01$ ,  $n = 234$ , Spearman's rho; 100 replications) (Fig. 3a). In all cases, compared with randomly selected subtrees of the same size from the reference tree, the correlations were the same as expected (Fig. 3a).

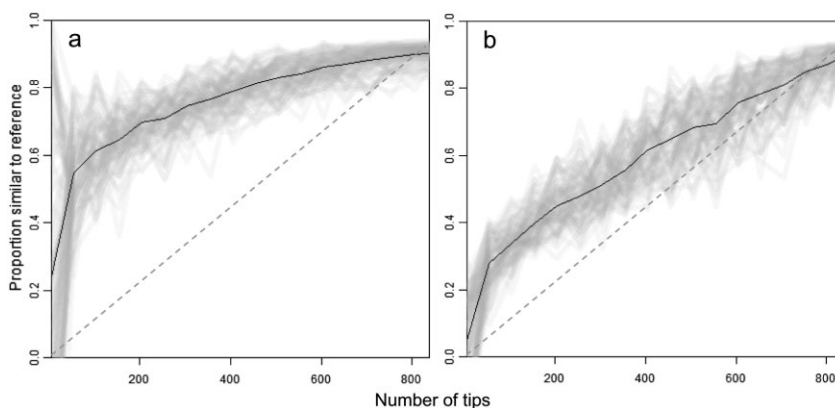
When imputing ED scores, accuracy in terms of correlation to the reference scores did not reduce as expected. Surprisingly, the 56 species for which there were expert-led scores (2010a) saw a large increase in correlation with the reference scores when they were used for estimating scores for all 838 species (from 0.63 to 0.79,  $n = 838$ , Spearman's rho; 100 replications), although the most accurate was the 2013 phylogeny (0.85,  $SD = 0.01$ ,  $n = 838$ , Spearman's rho; 100 replications) (Fig. 3b).

The number of 'reference' top 50 species were identified most efficiently using the workshop-derived ED scores which captured an average of 29 ( $SD = 2.24$ ; 100 replications) of the 'reference' top 50, compared with imputed scores from random 53-tip subtrees which identified ~10 species successfully. With the exception of the 2013 tree, the other methods captured more of the reference top 50 species than subsamples of the full tree (Fig. 4).

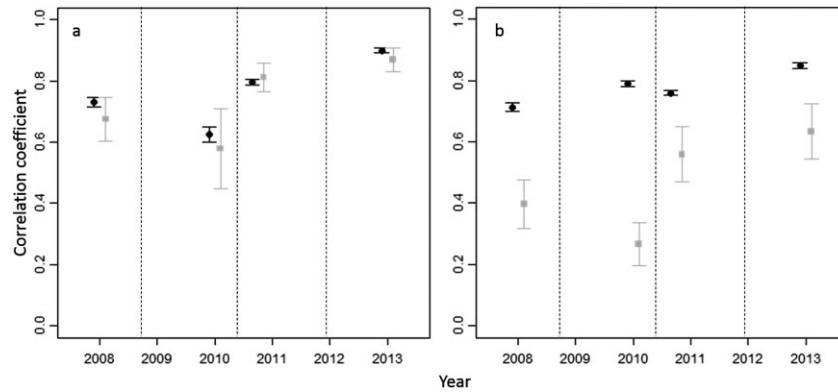
## Discussion

As expected, using the more complete and up-to-date phylogenies resulted in better prediction of reference scores. Surprisingly, the older phylogenies and scores derived from expert opinion also performed well. The expert opinion-led scores, when used for imputing score for missing species, was a close second in terms of prediction accuracy when compared with the most recent and largest tree, which had nearly 10 times the coverage. The key reason for this pattern is most likely a concerted attempt by the experts to consider species from every section of the tree rather than relying on those species that had already been sampled genetically (Rokas & Carroll, 2005).

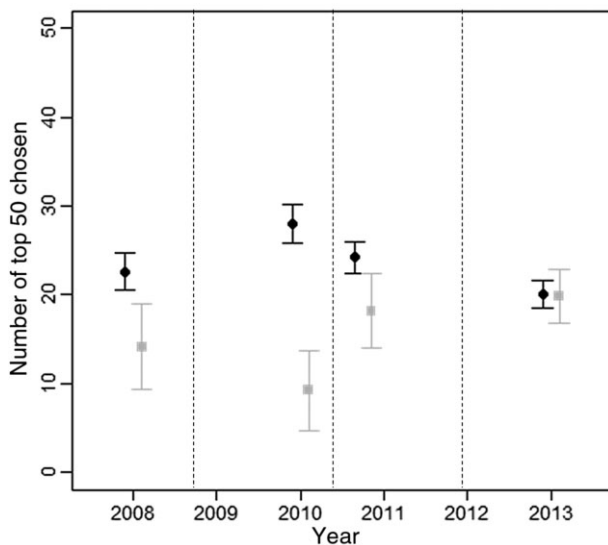
The analyses also show that our approach of imputing data by simple taxonomy-based extrapolation of ED using



**Figure 2** Correlation of evolutionary distinctiveness (ED) scores (a) from subsample  $n$  tip trees to scores calculated on the full 838-tip tree, from which subtrees are sampled. ED scores in (b) are imputed for the missing species using mean ED scores from closely related species. Dashed lines represent a hypothetical linear gain in accuracy as tree increases in size.



**Figure 3** Evolutionary distinctiveness (ED) scores (a) ( $n = 127, 56, 234$  and  $466$ , respectively) and imputed ED scores (b) ( $n = 838$  in all cases) (black circles) calculated on chronologically successive sets of incomplete, older phylogenetic trees correlated against reference ED scores that were calculated on a later full species-level supertree of corals. Grey squares mark correlation if ED scores were calculated on trees that were actually subsamples of the later tree but contain the same number of tips as the published phylogenies. Whiskers show 95% confidence intervals. For references see text.



**Figure 4** Successful identification of the 'reference' top 50 EDGE scoring species by chronologically successive sets of incomplete, older phylogenetic trees (black circles). Grey squares mark successful recognition if evolutionary distinctiveness scores were calculated on trees that were actually subsamples of the later tree but contain the same number of tips as the published phylogenies. Whiskers show 95% confidence intervals. For references see text.

incomplete phylogenetic knowledge is worthwhile. Interestingly, the score accuracy remains unchanged during this process or, in the case of expert-led scores, appears to increase. The reference tree incorporates much of the recent phylogenetic data that have emerged, yet the use of a coarse classification that is in conflict with these new findings, that is Veron (2000), produces ED scores that are very similar to the reference results. Given the ability of small trees to be extrapolated to full taxonomic coverage with limited cost, we recommend more species groups to be assessed for their

EDGE priorities (Isaac *et al.*, 2007) to prompt conservation action in the absence of data. We should certainly strive for reduced phylogenetic uncertainty and even fully resolved trees based on comprehensive DNA sequence data in order to carry out these analyses, but our results suggest that a fairly accurate picture of evolutionary distinctiveness can be obtained with smaller datasets. Furthermore, it may be possible to achieve even greater accuracy in the future and we recommend research that utilizes more sophisticated methods for imputing missing data, such as a Bayesian phylogenetic approach (Hadfield & Nakagawa, 2010; Wiens & Tiu, 2012).

Our results suggest that small well-resolved trees can also accurately estimate ED scores when computing scores for just those species sampled. Although it was the least complete tree, the 2008-derived tree appears to benefit from being a well-resolved multi-gene tree with very high support values. However, our reference dataset will need to be revised as further iterations of the phylogeny are developed and a better consensus is established.

The top 50 EDGE scleractinian species lists showed significant variation from one another and from the reference list. While the GE component of the EDGE score remained consistent across the five analyses, the ED scores were sensitive to variation in the analytical methods. For example, in constructing the supertree, those taxa for which only topological information was available were initially assigned a branch length of zero at the terminal tip (Huang, 2012). During the process of creating an ultrametric time tree, these tips in particular are greatly extended by the time-calibration process so that they coincides with the present. Similarly, other processes may have caused discrepancies in the creation of our reference ED dataset. For example, *Acropora echinata*, *Acropora speciosa* and *Favites abdita* are each nested within clades containing numerous species. From a purely topological point of view, they are relatively indistinct and when the time-calibration process is applied, their terminal edges are

not scaled upward as much as other species, lowering their ranks.

### Value of expert opinion

The workshop approach enables experts to contribute valuable, unpublished or field observations to complement the existing phylogenetic literature. This wealth of knowledge would be difficult for a single researcher to amass, but led to an improved accuracy of calculating EDGE scores by approximately 7% here, thus serving to validate the use of this method in setting conservation-priorities species lists. However, future assessors should consider whether this improved accuracy is cost and time effective. For example, the EDGE Coral Expert Workshop was devised to establish the 'top 10' species to which conservation efforts should be directed. The two working groups within the expert workshop initially generated a list of 12 potential EDGE species, with eight being independently and consistently selected by both groups, showing a high degree of consensus. Such an approach could be easily achieved using alternative methods of opinion gathering, that is online polls such as SurveyMonkey (SurveyMonkey Inc., Palo Alto, CA, USA; Wright, 2005).

The workshop was not limited to purely setting priorities. An additional benefit of an expert-led workshop approach is that species can be assessed against secondary criteria (i.e. functional redundancy, dependant species or level of conservation attention) thus, making the resulting species list more practical. Such criteria (Supporting Information Appendix S3) were adopted at the EDGE Coral Workshop to reduce the 12 species put forward by the independent groups to the finalized top 10 priority EDGE coral species (Supporting Information Appendix S4). However, the strong differences between the top 10 species lists generated by empirical datasets suggest that in early data-poor stages, such workshops should be directed towards identifying a larger number of priority species. In data-poor situations, focusing efforts on the top 10 species appears to risk a very high turnover of high-priority species with subsequent phylogenetic iterations. This has been the case with the EDGE Mammals List. However, there is always a trade-off between conservation urgency and developing knowledge.

With the 2008 tree, as with many smaller incomplete phylogenies, species-poor groups (i.e. those likely to have high ED scores) are often not represented by genetic data (e.g. *Ctenactis* or *Lithophyllon*). It may be that these species are geographically remote, scarce and/or have unusual ecologies (e.g. *Anomastrea irregularis* and *Horastrea indica*). The EDGE Coral Expert Workshop was effective at identifying the top 50 EDGE species because all species lineages were considered for listing, whether sequence data were available or not. Indeed, the estimated ED scores produced by the EDGE Coral Expert Workshop appeared to add value to the scores from the incomplete 2008 phylogeny, allowing for more accurate identification of the top 50 species.

In addition to identifying the top 10 EDGE coral species, the workshop also added value to setting conservation pri-

orities by identifying a list of ED species that are classified as data deficient (ED-DD species; Supporting Information Appendix S5) or not currently in a threatened category (Supporting Information Appendix S6) according to the IUCN Red List. These species should be monitored as they could become EDGE priority species in the future should their risk of extinction ever be assessed as threatened. Research into the extinction risk of these species is therefore identified as a high priority. Of these ED and ED-DD species (excluding those identified as the top 10 EDGE species), 77% were in the top 100 ED species of the reference list.

### Need for conservation action

Increasingly, it is being recognized by the scientific community that 'data-less and data-poor management are, under current circumstances, not just valid alternatives, but imperatives' (Johannes, 1998). This was one of the original motivations behind the EDGE Coral Programme and in spite of a lack of phylogenetic data for some groups, active conservation projects are now underway for seven of the top EDGE corals (*Catalaphyllia jardinei*, *Ctenella chagius*, *Dendrogyra cylindrus*, *Heliopungia actiniformis*, *Horastrea indica*, *Parasimplastrea sheppardi* and *Physogyra lichtensteini*) through the EDGE of Existence Programme.

While research is ongoing to enhance our phylogenetic knowledge of scleractinian corals and improve upon methodologies for phylogenetic-based conservation planning, the imminent and multiple threats to coral reefs necessitate urgent and decisive action. Therefore, despite the present incompleteness and uncertainty in phylogenetic inferences, an approach that uses the best science available to identify ED coral species is of great benefit to ongoing and future conservation initiatives. Here, we show that imputing data for missing species represents a promising development to allow phylogeny-led conservation programmes to assess and conserve those species groups we know little about, including the vast diversity of marine invertebrates that are even more data poor than corals. The dearth of EDGE assessments for invertebrates is indicative of our poor understanding of these animals (see Cardoso *et al.*, 2011). Fortunately, the approach we demonstrate here, including our handling of data-deficient species and tracking of ED scores with increasing phylogenetic information, can be applied more broadly across the metazoan tree of life.

### Acknowledgements

This research was supported by the Mohammed bin Zayed Species Conservation Fund and Zoological Society of London. We are grateful to Professor Charlie Veron, Professor Jonathan Baillie, Dr Arjan Gittenberger, Dr Rosemarie Baron-Szabo, Helen Meredith, Dr Francesca Benzoni, Dr Sam Turvey, Carly Waterman and Dr Simon Harding for their input and insights into the EDGE Coral Workshop process. We also acknowledge the contribution of the many coral experts who were unable to attend the meeting but formed a wider consultation group.

## References

- Arrigoni, R., Stefani, F., Pichon, M., Galli, P. & Benzoni, F. (2012). Molecular phylogeny of the Robust clade (Faviidae, Mussidae, Merulinidae, and Pectiniidae): an Indian Ocean perspective. *Mol. Phylogenet. Evol.* **65**, 183–193.
- Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O.U., Swartz, B., Quental, T.B., Marshall, C., McGuire, J.L., Lindsey, E.L. & Maguire, K.C. (2011). Has the earth's sixth mass extinction already arrived? *Nature* **471**, 51–57.
- Baum, B.R. (1992). Combining trees as a way of combining data sets for phylogenetic inference, and the desirability of combining gene trees. *Taxon* **41**, 3–10.
- Benzoni, F., Stefani, F., Stolarski, J., Pichon, M., Mitta, G. & Galli, P. (2007). Debating phylogenetic relationships of the scleractinian *Psammocora*: molecular and morphological evidences. *Contrib. Zool.* **76**, 35–54.
- Bibby, C., Collar, N.J., Crosby, M., Heath, M., Imboden, C., Johnson, T., Long, A., Stattersfield, A. & Thirgood, S. (1992). *Putting biodiversity on the map: priority areas for global conservation*. Cambridge: International Council for Bird Preservation.
- Brooks, T.M., Mittermeier, R.A., da Fonseca, G.A.B., Gerlach, J., Hoffmann, M., Lamoreux, J.F., Mittermeier, C.G., Pilgrim, J.D. & Rodrigues, A.S.L. (2006). Global biodiversity conservation priorities. *Science* **313**, 58–61.
- Budd, A.F., Romano, S.L., Smith, N.D. & Barbeitos, M.S. (2010). Rethinking the phylogeny of scleractinian corals: a review of morphological and molecular data. *Integr. Comp. Biol.* **50**, 411–427.
- Cadotte, M.W., Cardinale, B.J. & Oakley, T.H. (2008). Evolutionary history and the effect of biodiversity on plant productivity. *Proc. Natl. Acad. Sci. USA* **105**, 17013–17017.
- Cadotte, M.W., Cavender-Bares, J., Tilman, D. & Oakley, T.H. (2009). Using phylogenetic, functional and trait diversity to understand patterns of plant community productivity. *PLoS ONE* **4**, e5695.
- Cardoso, P., Erwin, T.L., Borges, P.A.V. & New, T.R. (2011). The seven impediments in invertebrate conservation and how to overcome them. *Biol. Conserv.* **144**, 2647–2655.
- Carpenter, K.E., Abrar, M., Aeby, G., Aronson, R.B., Banks, S., Bruckner, A., Chiriboga, A., Cortés, J., Delbeek, J.C., Devantier, L., Edgar, G.J., Edwards, A.J., Fenner, D., Guzmán, H.M., Hoeksema, B.W., Hodgson, G., Johan, O., Licuanan, W.Y., Livingstone, S.R., Lovell, E.R., Moore, J.A., Obura, D.O., Ochavillo, D., Polidoro, B.A., Precht, W.F., Quibilan, M.C., Reboton, C., Richards, Z.T., Rogers, A.D., Sanciangco, J., Sheppard, A., Sheppard, C., Smith, J., Stuart, S., Turak, E., Veron, J.E.N., Wallace, C., Weil, E. & Wood, E. (2008). One-third of reef-building corals face elevated extinction risk from climate change and local impacts. *Science* **321**, 560–563.
- Castresana, J. (2007). Topological variation in single-gene phylogenetic trees. *Genome Biol.* **8**, 216.
- Collen, B., Turvey, S.T., Waterman, C., Meredith, H.M.R., Kuhn, T.S., Baillie, J.E.M. & Isaac, N.J.B. (2011). Investing in evolutionary history: implementing a phylogenetic approach for mammal conservation. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **366**, 2611–2622.
- Cooper, A. & Fortey, R. (1998). Evolutionary explosions and the phylogenetic fuse. *Trends Ecol. Evol.* **13**, 151–156.
- Crozier, R.H. (1992). Genetic diversity and the agony of choice. *Biol. Conserv.* **61**, 11–15.
- De Vos, J.M., Joppa, L.N., Gittleman, J.L., Stephens, P.R. & Pimm, S.L. (2014). Estimating the normal background rate of species extinction. *Conserv. Biol.* (Online DOI: 10.1111/cobi.12380).
- Donner, S.D., Skirving, W.J., Little, C.M., Oppenheimer, M. & Hoegh-Guldberg, O. (2005). Global assessment of coral bleaching and required rates of adaptation under climate change. *Glob. Change Biol.* **11**, 2251–2265.
- Dove, S.G., Kline, D.I., Pantos, O., Angly, F.E., Tyson, G.W. & Hoegh-Guldberg, O. (2013). Future reef decalcification under a business-as-usual CO<sub>2</sub> emission scenario. *Proc. Natl. Acad. Sci. USA* **110**, 15342–15347. (Online DOI: 10.1073/pnas.1302701110).
- Drummond, A.J. & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214.
- Faith, D.P. (1992). Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* **61**, 1–10.
- Faith, D.P. (1996). Conservation priorities and phylogenetic pattern. *Conserv. Biol.* **10**, 1286–1289.
- Faith, D.P. (2007). Phylogeny and conservation. *Syst. Biol.* **56**, 690–694.
- Fukami, H., Chen, C.A., Budd, A.F., Collins, A., Wallace, C., Chuang, Y.Y., Chen, C., Dai, C.F., Iwao, K., Sheppard, C. & Knowlton, N. (2008). Mitochondrial and nuclear genes suggest that stony corals are monophyletic but most families of stony corals are not (Order Scleractinia, Class Anthozoa, Phylum Cnidaria). *PLoS ONE* **3**, e3222.
- Gittenberger, A., Reijen, B.T. & Hoeksema, B.W. (2011). A molecularly based phylogeny reconstruction of mushroom corals (Scleractinia: Fungiidae) with taxonomic consequences and evolutionary implications for life history traits. *Contrib. Zool.* **80**, 107–132.
- Hadfield, J.D. & Nakagawa, S. (2010). General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *J. Evol. Biol.* **23**, 494–508.
- Hellberg, M.E. (2006). No variation and low synonymous substitution rates in coral mtDNA despite high nuclear variation. *BMC Evol. Biol.* **6**, 24.



- Hoegh-Guldberg, O., Mumby, P.J., Hooten, A.J., Steneck, R.S., Greenfield, P., Gomez, E., Harvell, C.D., Sale, P.F., Edwards, A.J., Caldeira, K., Knowlton, N., Eakin, C.M., Iglesias-Prieto, R., Muthiga, N., Bradbury, R.H., Dubi, A. & Hatziolos, M.E. (2007). Coral reefs under rapid climate change and ocean acidification. *Science* **318**, 1737–1742.
- Hoffmann, M., Brooks, T., da Fonseca, G., Gascon, C., Hawkins, A., James, R., Langhammer, P., Mittermeier, R., Pilgrim, J., Rodrigues, A. & Silva, J. (2008). Conservation planning and the IUCN Red List. *Endanger. Species Res.* **6**, 113–125.
- Holland, B., Conner, G., Huber, K. & Moulton, V. (2006). Imputing supertrees and supernetworks from quartets. *Syst. Biol.* **56**, 57–67.
- Huang, D. (2012). Threatened reef corals of the world. *PLoS ONE* **7**, e34459.
- Huang, D. & Roy, K. (2013). Anthropogenic extinction threats and future loss of evolutionary history in reef corals. *Ecol. Evol.* **3**, 1184–1193.
- Huang, D., Meier, R., Todd, P.A. & Chou, L.M. (2008). Slow mitochondrial COI sequence evolution at the base of the metazoan tree and its implications for DNA barcoding. *J. Mol. Evol.* **66**, 167–174.
- Huang, D., Meier, R., Todd, P.A. & Chou, L.M. (2009). More evidence for pervasive paraphyly in scleractinian corals: systematic study of Southeast Asian Faviidae (Cnidaria; Scleractinia) based on molecular and morphological data. *Mol. Phylogenet. Evol.* **50**, 102–116.
- Huelsenbeck, J.P. & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755.
- International Union for the Conservation of Nature (IUCN). (2013). *IUCN red list of threatened species*. Available at <http://www.redlist.org> (accessed January 2014).
- Isaac, N.J.B., Turvey, S.T., Collen, B., Waterman, C. & Baillie, J.E.M. (2007). Mammals on the EDGE: conservation priorities based on threat and phylogeny. *PLoS ONE* **2**, e296. (Online DOI: 10.1371/journal.pone.0000296).
- Isaac, N.J.B., Redding, D.W., Meredith, H.M.R. & Safi, K. (2012). Phylogenetically-informed conservation priorities for amphibian conservation. *PLoS ONE* **7**, e43912.
- Jetz, W., Thomas, G.H., Joy, J.B., Redding, D.W., Hartmann, K. & Mooers, A.O. (2014). Global distribution and conservation of evolutionary distinctness in birds. *Curr. Biol.* **24**, 919–930.
- Johannes, R.E. (1998). The case for data-less marine resource management: examples from tropical nearshore finfisheries. *Trends Ecol. Evol.* **13**, 243–246.
- Kennedy, E.V., Perry, C.T., Halloran, P.R., Iglesias-Prieto, R., Schönberg, C.H.L., Max Wisshak, M., Form, A.U., Carricart-Ganivet, J.P., Fine, C., Eakin, M. & Mumby, P.J. (2013). Avoiding coral reef functional collapse requires local and global action. *Curr. Biol.* **23**, 912–918.
- Kerr, A.M. (2005). Molecular and morphological supertree of stony corals (Anthozoa: Scleractinia) using matrix representation parsimony. *Biol. Rev. Camb. Philos. Soc.* **80**, 543–558.
- Kitahara, M., Cairns, S.D., Stolarski, J., Blair, D. & Miller, D.J. (2010). A comprehensive phylogenetic analysis of the Scleractinia (Cnidaria, Anthozoa) based on mitochondrial COI sequence data. *PLoS ONE* **5**, e11490.
- Kuhn, T.S., Mooers, A. & Thomas, G.H. (2011). A simple polytomy resolver for dated phylogenies. *Methods Ecol. Evol.* **2**, 427–436.
- Little, R.J.A. & Rubin, D.B. (1987). *Statistical analysis with missing data*. New York: Wiley.
- Mace, G.M., Collar, N.J., Gaston, K.J., Hilton-Taylor, C., Akçakaya, H.R., Leader-Williams, N., Milner-Gulland, E.J. & Stuart, S.N. (2008). Quantification of extinction risk: IUCN's system for classifying threatened species. *Conserv. Biol.* **22**, 1424–1442.
- Margules, C.R. & Sarkar, S. (2006). *Systematic conservation planning*. New York: Cambridge University Press.
- Mittermeier, R.A., Myers, N., Thomsen, J.B., Gustavo, A.B. & Olivieri, S. (1998). Biodiversity hotspots and major tropical wilderness areas: approaches to setting conservation priorities. *Conserv. Biol.* **12**, 516–520.
- Mooers, A.O., Faith, D.P. & Maddison, W.P. (2008). Converting endangered species categories to probabilities of extinction for phylogenetic conservation prioritization. *PLoS ONE* **3**, e3700.
- Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B. & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858.
- Nunes, F., Fukami, H., Vollmer, S.V., Norris, R.D. & Knowlton, N. (2008). Re-evaluation of the systematics of the endemic corals of Brazil by molecular data. *Coral Reefs* **27**, 423–432.
- O'Connor, C., Marvier, M. & Kareiva, P. (2003). Biological versus social, economic and political priority-setting in conservation. *Ecol. Lett.* **6**, 706–711.
- Olson, D.M. & Dinerstein, E. (1998). The global conserving representation approach earth's most biologically valuable ecoregions. *Conserv. Biol.* **12**, 502–515.
- Orme, D., Freckleton, R., Thomas, G.H., Petzoldt, T., Fritz, S., Isaac, N.J.B. & Pearse, W. (2012). caper: Comparative Analyses of Phylogenetics and Evolution in R. R Packag. version 0.5. Available at <http://caper.r-forge.r-project.org/> (accessed May 2014).
- Pimm, S.L., Raven, R., Peterson, A., Şekerciöglu, C.H. & Ehrlich, P.R. (2006). Human impacts on the rates of recent, present, and future bird extinctions. *PNAS* **103**, 10941–10946.
- Possingham, H.P. & Wilson, K.A. (2005). Biodiversity: turning up the heat on hotspots. *Nature* **436**, 919–920.
- Pressey, R.L., Cabeza, M., Watts, M.E., Cowling, R.M. & Wilson, K.A. (2007). Conservation planning in a changing world. *Trends Ecol. Evol.* **22**, 583–592.

- R Core Team. (2013). R: a language and environment for statistical computing. *R Found. Stat. Comput. Vienna, Austria*. Available at <http://www.r-project.org/> (accessed April 2014).
- Ragan, M.A. (1992). Phylogenetic inference based on matrix representation of trees'. *Mol. Phylogenet. Evol.* **1**, 53–58.
- Redding, D.W. & Mooers, A.Ø. (2006). Incorporating evolutionary measures into conservation prioritization. *Conserv. Biol.* **20**, 1670–1678.
- Redding, D.W., Hartmann, K., Mimoto, A., Bokal, D., DeVos, M. & Mooers, A.Ø. (2008). Evolutionarily distinctive species often capture more phylogenetic diversity than expected. *J. Theor. Biol.* **251**, 606–615.
- Redding, D.W., Dewolff, C.V. & Mooers, A.Ø. (2010). Evolutionary distinctiveness, threat status, and ecological oddity in primates. *Conserv. Biol.* **24**, 1052–1058.
- Rodrigues, A.S.L., Pilgrim, J.D., Lamoreux, J.F., Hoffmann, M. & Brooks, T.M. (2006). The value of the IUCN Red List for conservation. *Trends Ecol. Evol.* **21**, 71–76.
- Rokas, A. & Carroll, S.B. (2005). More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy. *Mol. Biol. Evol.* **22**, 1337–1344.
- Ronquist, F. & Huelsenbeck, J.P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574.
- Safi, K., Armour-Marshall, K., Baillie, J.E.M. & Isaac, N.J.B. (2013). Global patterns of evolutionary distinct and globally endangered amphibians and mammals. *PLoS ONE* **8**, e63582.
- Selig, E.R., Turner, W.R., Troeng, S., Wallace, B.P., Halpern, B.S., Kaschner, K., Lascelles, B.G., Carpenter, K.E. & Mittermeier, R.A. (2014). Global priorities for marine biodiversity conservation. *PLoS ONE* **9**, e82898. (Online DOI: 10.1371/journal.pone.0082898).
- Shearer, T.L., van Oppen, M.J.H., Romano, S.L. & Wörheide, G. (2002). Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). *Mol. Ecol.* **11**, 2475–2487.
- Sheppard, C.R.C. (2003). Predicted recurrences of mass coral mortality in the Indian Ocean. *Nature* **425**, 294–297.
- Simpson, C., Kiessling, W., Mewis, H., Baron-Szabo, R.C. & Müller, J. (2011). Evolutionary diversification of reef corals: a comparison of the molecular and fossil records. *Evolution* **65**, 3274–3284.
- Stolarski, J., Kitahara, M.V., Miller, D.J., Cairns, S.D., Mazur, M. & Meibom, A. (2011). The ancient evolutionary origins of Scleractinia revealed by azooxanthellate corals. *BMC Evol. Biol.* **11**, 316. (Online DOI: 10.1186/1471-2148-11-316).
- Swenson, N.G. (2013). Phylogenetic imputation of plant functional trait databases. *Ecography* **37**, 105–110.
- Treblico, R., Halpern, B.S., Flemming, J.M., Field, C., Blanchard, W. & Worm, B. (2011). Mapping species richness and human impact drivers to inform global pelagic conservation prioritisation. *Biol. Conserv.* **144**, 1758–1766.
- Vane-Wright, R.I., Humphries, C.J. & Williams, P.H. (1991). What to protect? Systematics and the agony of choice. *Biol. Conserv.* **55**, 235–254.
- Veron, J.E.N. (2000). *Corals of the world (3 vols)*. Townsville: Australian Institute of Marine Science.
- Veron, J.E.N., Hoegh-Guldberg, O., Lenton, T.M., Lough, J.M., Obura, D.O., Pearce-Kelly, P., Sheppard, C.R.C., Spalding, M., Stafford-Smith, M.G. & Rogers, A.D. (2009). The coral reef crisis: the critical importance of <350 ppm CO<sub>2</sub>. *Mar. Pollut. Bull.* **58**, 1428–1436.
- Waldron, A., Mooers, A.O., Miller, D.C., Nibbelink, N., Redding, D., Kuhn, T.S., Roberts, J.T. & Gittleman, J.L. (2013). Targeting global conservation funding to limit immediate biodiversity declines. *Proc. Natl. Acad. Sci. USA* **110**, 12144–12148.
- Wiens, J.J. & Tiu, J. (2012). Highly incomplete taxa can rescue phylogenetic analyses from the negative impacts of limited taxon sampling. *PLoS ONE* **7**, e42925.
- Wilson, K.A., McBride, M.F., Bode, M. & Possingham, H.P. (2006). Prioritizing global conservation efforts. *Nature* **440**, 337–340.
- Witting, L. & Loeschcke, V. (1995). The optimization of biodiversity conservation. *Biol. Conserv.* **71**, 205–207.
- Wright, K.B. (2005). Researching internet-based populations: advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *J. Comput. Med. Commun.* **10**, 00. Available at <http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2005.tb00259.x/full> (accessed on June 2014).
- WWF (2014). *Living Planet Report 2014: species and spaces, people and places*. McLellan, R., Iyengar, L., Jeffries, B. & Oerlemans, N. (Eds). Gland, Switzerland: World Wildlife Fund.

## Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

**Appendix S1.** Species shortlisted for consideration for 2010 EDGE Coral Workshop.

**Appendix S2.** ED scoring criteria for 2010 EDGE Coral Workshop.

**Appendix S3.** Secondary criteria for 2010 EDGE Coral Workshop.

**Appendix S4.** 2010 EDGE Coral Workshop Top 10 species.

**Appendix S5.** Summary of species identified as ED-DD species at the 2010 EDGE Coral Workshop.

**Appendix S6.** Summary of species identified as ED species at the 2010 EDGE Coral Workshop but not included in the EDGE list with IUCN priority listings.