

site-selection algorithms to assess whether a network of areas selected to maximize the number of surrogate taxa captures a greater number of target taxa than expected by chance; Table 1). The sample size of each surrogacy test (n) was also recorded to account for different sampling intensities across studies.

From each surrogacy test, whether or not the null hypothesis was rejected (based on the authors' arbitrarily chosen threshold), and the coefficient of determination (R^2) of surrogate predictive power were collated. Overall surrogate effectiveness was then defined using both the probability of concluding that a surrogate performed better than random (P) and the surrogate predictive power (R^2). Ideally, surrogate effectiveness would be assessed as the likelihood of surrogate non-randomness based on bias-corrected maximum likelihoods for multi-model comparisons [22] or secondarily, as the probability of making a Type I error when concluding that surrogate predictions performed better than random. Such approaches could not be adopted here because most surrogacy studies have only reported an arbitrary probability threshold (e.g., $P < 0.05$) when testing the null hypothesis that the surrogate did not perform better than random.

A central assumption of meta-analyses is that the literature reviewed is not subject to publication bias, which can arise if the probability an article is published depends on the strength and direction of its results [19]. We tested for publication bias using funnel plots of sample size against effect size, i.e., the number of tests in each 0.05-class of R^2 (not shown; [23]) combined with a rank correlation test between standardized sample size and effect size.

Elicitation of priors using expert knowledge

Eight international experts in marine biological surrogacy answered an online survey (<http://www.adelaide.edu.au/environment/mbp/survey/02.html>) on the effectiveness of biological surrogates in marine ecosystems. They were asked to estimate the likelihood (P) of a surrogate being effective in predicting a target, and the proportion of variance explained (R^2) by the

surrogate for the different habitats, scales, types of surrogates and statistical approaches defined above. In each situation, experts were asked to classify P and R^2 into one of five categories (0.0–0.2; 0.2–0.4; 0.4–0.6; 0.6–0.8; 0.8–1). Additionally, they were asked to indicate their confidence in the classification using the same categories. We translated these classes into a categorical estimate from 'very low/small' to 'very high/large' to facilitate interpretation by the experts [20]. Their type and level of expertise were assessed using questions on their research activities and proportion of time allocated to the study of marine biological surrogates. The effect of their level of expertise and how it was acquired on their confidence when answering these questions was investigated with a multivariate analysis of variance using a Bray-Curtis distance matrix and 100 permutations (NPMANOVA; [24]). The statistical units were experts, and multivariate responses were their mean confidence score when answering questions relating to the *Habitat*, *Scale*, *Type*, or *Stats* factors. Predictors were the descriptors of their expertise (Table S1).

Bayesian model fitting

Bayesian hierarchical (i.e., multilevel) models of surrogate effectiveness, successively defined as P and R^2 , were implemented to assess the influence of the factors defined above on surrogate effectiveness. For each response variable, covariates included each of the factors *Habitat*, *Scale*, *Type*, *Stats* coded as dummy variables, or n , in a separate model. The hierarchical term *Study* was added to account for the non-independence of multiple tests within the same study. The resulting model formulation is given by:

$$\text{logit}(y_{ij}) = \sum_k \beta_k X_{ijk} + \epsilon_{0j}$$

$$\epsilon_{0j} \sim N(0, \tau_j)$$

Table 1. Statistical methods and biodiversity metrics used in marine biological surrogacy studies.

Method	Statistical index	Biodiversity metrics	Description of biodiversity metric	Test ID (examples)
■ Congruence of univariate biodiversity metrics	■ Spearman's rho or Pearson's r	■ Taxonomic richness	■ Number of taxa (e.g., species, genera)	1–3, 60–65, 98–101
		■ Numerical Rarity	■ Number of rare species (e.g., $n \leq 2$)	57, 59
		■ Endemicity	■ Number of endemic species (e.g., based on range extent)	102
		■ Abundance	■ Number of individuals (per unit area)	74, 218, 220–222
		■ Biomass	■ Mass per unit area	73, 75
■ Congruence of multivariate biodiversity metrics	■ Spearman's rho or Pearson's r between surrogate and target	■ Incidence matrix	■ Presence or absence of taxa (columns) at the different sites (rows)	4, 9–23
		■ Bray-Curtis distance matrices	■ Abundances of taxa (columns) at the different sites (rows)	76–97, 103–108
■ Representation	■ Site-selection algorithms	■ Taxonomic richness	■ Number of taxa (e.g., species, genera)	5, 7, 25–55
		■ Numerical Rarity	■ Number of rare species	6, 8
		■ Occurrence of assemblages	■ Taxa grouped according to similarity in distribution	149–162

Methods used are grouped into three categories: congruence of univariate biodiversity metrics assesses whether surrogate biodiversity is spatially correlated with target biodiversity; congruence of multivariate biodiversity metrics evaluates whether pairs of sites showing the highest similarity in surrogate assemblages also show the highest similarity in target assemblages, and; representation uses site-selection algorithms to assess whether a network of areas selected to maximize the number of surrogate taxa also maximises the number of target taxa and whether this number is greater than expected by chance. Test ID refers to Table S2.

doi:10.1371/journal.pone.0020141.t001

where y_{ij} is the surrogate effectiveness for each test i of study j , β_k is the coefficient associated with the k^{th} (dummy) covariate X_{ijk} , ϵ_{oj} is the effect for the j^{th} study and τ_j is gamma-distributed ($a=0.001$, $b=0.001$). The binary P response variable was modeled using a binomial density function. The continuous R^2 response variable, which is restricted to the interval $[0, 1]$, was modeled using a beta density function (see code in Text S1; [25]). The logit link function was used for both response variables. Both uninformative priors $\beta_k \sim N(0, 1000)$ and informative priors $\beta_k \sim N(\mu_k, \sigma_k^2)$, where μ_k and σ_k^2 were estimated based on expert opinion, were considered. Therefore, for each response variable and each covariate, three models were considered with (i) covariate effect only and uninformative priors (ii) covariate and hierarchical effects and uninformative priors and (iii) covariate and hierarchical effects and informative priors. Model performance was assessed using the deviance information criterion (DIC) [26]. Model parameters were estimated using Markov chain Monte Carlo (MCMC) and Gibbs sampling. To ensure model convergence, models were run for 50,000 iterations with a 2,000 iteration burn-in period and every 12th observation was retained to control for any potential autocorrelation; the remaining 4000 values of each parameter were retained to generate the posterior distributions of the parameters. Model fitting was done using WinBUGS 1.4 [27] and the R2WinBUGS [28] package for R 2.9.2 [29].

Results

Literature review and meta-data compilation

Peer-reviewed and published literature on biological surrogacy in marine ecosystems included 20 studies presenting 264 biological surrogacy tests (Table S2; Text S2). Of these 264 tests, 138 were for soft bottom habitats (all in temperate regions), 71 were for temperate reefs and 55 for coral reefs (see Table 2 for cross-factor tabulations). These surrogacy tests were distributed globally, with 10 tests from the Arctic, 114 from Australia, 55 from the Indo-Pacific (including Indonesia, Madagascar, the Philippines and Papua New Guinea), 59 from the Mediterranean Sea and 26 from northern Europe. For higher-taxa surrogates, there were a variable number of taxonomic steps between the surrogate and the target,

and the surrogate predictive power decreased as the number of taxonomic steps between the surrogate and the target increased (Table 3). We found no evidence of any publication bias in the surrogate predictive power: i.e., no consistent trend between the number of published articles and surrogate effectiveness (Spearman's $\rho = -0.20$; $P = 0.38$).

Elicitation of priors using expert knowledge

Most experts answered all questions within 15 to 30 minutes (Table S1) and left less than 10% of the questions (24 of 288) unanswered. Expert confidence averaged $58 \pm 13\%$ (mean \pm standard deviation), corresponding to 'fairly confident' according to the survey terminology. Mean expert confidence was influenced by their experience in statistical analysis only (NPMANOVA $R^2 = 0.30$; $P = 0.040$).

Experts' ranking of P was the highest in soft bottom habitats, at a 10–100 km spatial scale, using representation-based statistical methods and a higher-taxa surrogate (Figure 2: circles). Experts' ranking of R^2 was in agreement with that of P , although differences among factor levels were less pronounced (Figure 3: circles).

Bayesian model fitting

The *Type* model of the P response variable was top-ranked according to the deviance information criterion (DIC), with higher- taxa surrogates, and subset-taxa surrogates to a lesser extent, both predicting higher P than cross-taxa surrogates (odds ratio = 60.1 and 10.0, respectively; Table 4; Figure 4). The type of habitat best explained R^2 , with both soft bottoms and temperate reefs performing better than tropical reef (odds ratio = 7.3 and 6.4, respectively; Table 4; Figure 5). All factors were important predictors of P , but only the type of habitat and the sample size models were ranked higher than the null model for the R^2 response (Table 4). For both response variables, hierarchical models that incorporated a *Study* effect accounting for the non-independence of tests within the same study ranked higher than the covariate-effects-only models (Table 4).

Models incorporating an informative prior were ranked higher than those incorporating an uninformative prior for the *Method*

Table 2. Cross-tabulations of the number of tests for each combination of factor levels.

		Habitat			Type			Scale			Method		
		TR	TE	SO	CT	HT	ST	L	M	S	UC	MC	RP
Habitat	TR	55			55	0	0	2	53	0	5	16	34
	TE		71		20	36	15	29	27	15	24	39	8
	SO			138	27	92	19	16	80	42	32	84	22
Type	CT				102			22	65	15	21	44	37
	HT					128		14	77	37	20	86	22
	ST						34	11	18	5	20	9	5
Scale	L							47			19	28	0
	M								160		24	80	56
	S									57	18	31	8
Method	UC										61		
	MC											139	
	RP												64

The diagonal indicates the total number of tests in each factor level including, for *Habitat*, TR: tropical reefs, TE: temperate reefs, SO: soft bottoms; for *Type*, CT: cross-taxa surrogate, HT: higher-taxa surrogate, ST: subset-taxa surrogate; for *Scale*, L: >100 km, M: 10–100 km, S: <10 km; and for *Method*, UC: univariate congruence, MC: multivariate congruence, RP: representation.

doi:10.1371/journal.pone.0020141.t002

Table 3. Higher-taxa surrogate predictive power (R^2) as a function of the number of taxonomic steps between the surrogate and the target.

Nb steps	Surrogate	Target	R^2 mean	R^2 sd	n	mean R^2	n_{tot}
1	class	order	0.19	0.08	5	0.61	36
	order	family	0.47	0.06	6		
	family	genus	0.87	0.08	6		
	genus	species	0.91	0.09	19		
2	class	family	0.11	0.09	5	0.43	34
	order	genus	0.39	0.1	6		
	family	species	0.78	0.19	23		
3	class	genus	0.11	0.09	5	0.29	23
	order	species	0.48	0.25	18		
4	class	species	0.25	0.18	10	0.25	10

With sd = standard deviation, n = number of tests, n_{tot} = total number of tests. doi:10.1371/journal.pone.0020141.t003

model predicting P (Figure 2: diamonds; Table 4), and for the *Type*, *Method* and *Scale* models predicting R^2 (Figure 3: diamonds; Table 4). However, differences in DIC were generally small (i.e., $\Delta DIC < 1$), except for the *Scale* and *Type* models of P , indicating that models with uninformative or informative priors provided an approximately equal description of the data and were essentially undistinguishable. For the *Scale* and *Type* models of P , models with uninformative priors performed better than those with informative priors (Figure 2: squares). In the first case, priors based on expert opinion did not capture the pattern revealed by the literature, i.e., a decrease in surrogate effectiveness as spatial scale increases. In the second case, the ranking of factor levels according to expert opinion was in agreement with the meta-data; however, informative priors did not improve the model, as a consequence of their diffuse distribution and a precise estimate of the likelihood based on the meta-data only. The same situation was observed for the *Habitat* model of R^2 (Figure 3: squares).

Convergence was successfully achieved for all models, although we could not fit any model that included interactions because of missing cross-factor combinations. Models including additive effects of all covariates could not be fitted either; adding all terms aliased the random effect as a result of the low level of replication of the different cross-factor combinations across studies.

Discussion

Drivers of surrogate effectiveness

As a first attempt to collate information on biological surrogate effectiveness in marine ecosystems, our study highlighted overall that in most situations, surrogate effectiveness was typically lower than generally assumed. Even with a relatively low expectation of the relationship between surrogate and target biodiversity (i.e., non-random), in most situations more than one third of all studies found no evidence for such a relationship. Moreover, even when there was strong evidence for a relationship between surrogate and target biodiversity, the predictability of such a relationship was nevertheless often weak (e.g., at a spatial scale of 10–100 km; $P = 0.66 \pm 0.10$ and $R^2 = 0.35 \pm 0.05$). This clearly highlights the need to understand why some surrogates might not be appropriate in certain situations and to formulate precise recommendations for

a more reliable use of biological surrogates in future studies of marine biodiversity.

By combining expert knowledge and published literature on surrogate effectiveness, we showed that the spatial scale, habitat, type of surrogate and method used for its construction all influenced surrogate effectiveness, according to at least one of the two measures of effectiveness considered. The type of surrogate used was the strongest determinant of P , with higher-taxa surrogates predicting higher P than all other types. This greater effectiveness of higher-taxa surrogates might be intuitively expected because of the hierarchical relationship among taxonomic levels [30] where the probability of observing a high number of genera increases with the number of species observed. However even though expected, this result is demonstrated here for the first time and warrants further development of higher-taxa surrogates, once one guards against a number of potential pitfalls. First, the rate of spatial variation in biodiversity metrics (β diversity) declines with decreasing taxonomic resolution [31], so inter-site differences in species richness or composition might not be detected by such higher-taxonomic level surrogates. Secondly, we advise caution when comparing the effectiveness of higher-taxa surrogates across taxonomic groups. Indeed, subdivisions of taxa and their rationale for classification into various taxonomic levels is inconsistent across groups because current taxonomic classifications result from a heterogeneous mixture of various historical and contemporary views [32]. Therefore, we contend that higher-taxa approaches can provide valuable surrogates only at a scale where they reflect species-level patterns of β diversity, and as long as the inherent uncertainty of taxonomic classifications tempers conclusions.

Our results indicate that surrogates based on representation were less effective than those based on spatial congruence of univariate or multivariate biodiversity metrics. Representation-based methods, which consist of selecting a set of sites based on a surrogate and summing the representation of the target within the selected set, suffer from a number of flaws previously highlighted in studies of surrogacy in terrestrial ecosystems. Site-selection algorithms return one solution for reserve design from a large number of potentials which might not be optimal because they cannot assess the overall pattern of representation [33]. Moreover, representation-based surrogates are designed to perform better than the random addition of sites, but they are rarely compared to the optimum selection of sites, derived by selecting sites on the basis of the target instead of the surrogate – an approach that has been developed in terrestrial ecosystems [34]. Such methods have yet to be applied in marine ecosystems but appear worthy of exploration in this realm. Lastly, representation-based methods can and should be used with biodiversity metrics other than just taxonomic richness. The number of taxa observed at a given place depends strongly on sampling effort [35], and this might influence the relationship between surrogate and target species richness.

Surrogate effectiveness also decreased with increasing spatial scale. Most studies conducted at spatial scales > 100 km assumed homogeneity of the study system at finer scales. Such an assumption might be incorrect if different biogeographic sub-regions or distinct evolutionary histories are included [36,37], or if different combinations of habitats are represented among areas. Indeed, surrogate performance varies among areas displaying regional variation of species distributions [8,38] and patterns of 'local endemism' (i.e. fine-scale patterns in species-habitat associations; [37]). Despite some evidence of the importance of spatial scale on surrogate effectiveness, such effects are still largely unknown.

Surrogate effectiveness varied among habitats and was lowest for tropical coral reefs. The high biodiversity and functional

P

Parameter estimate

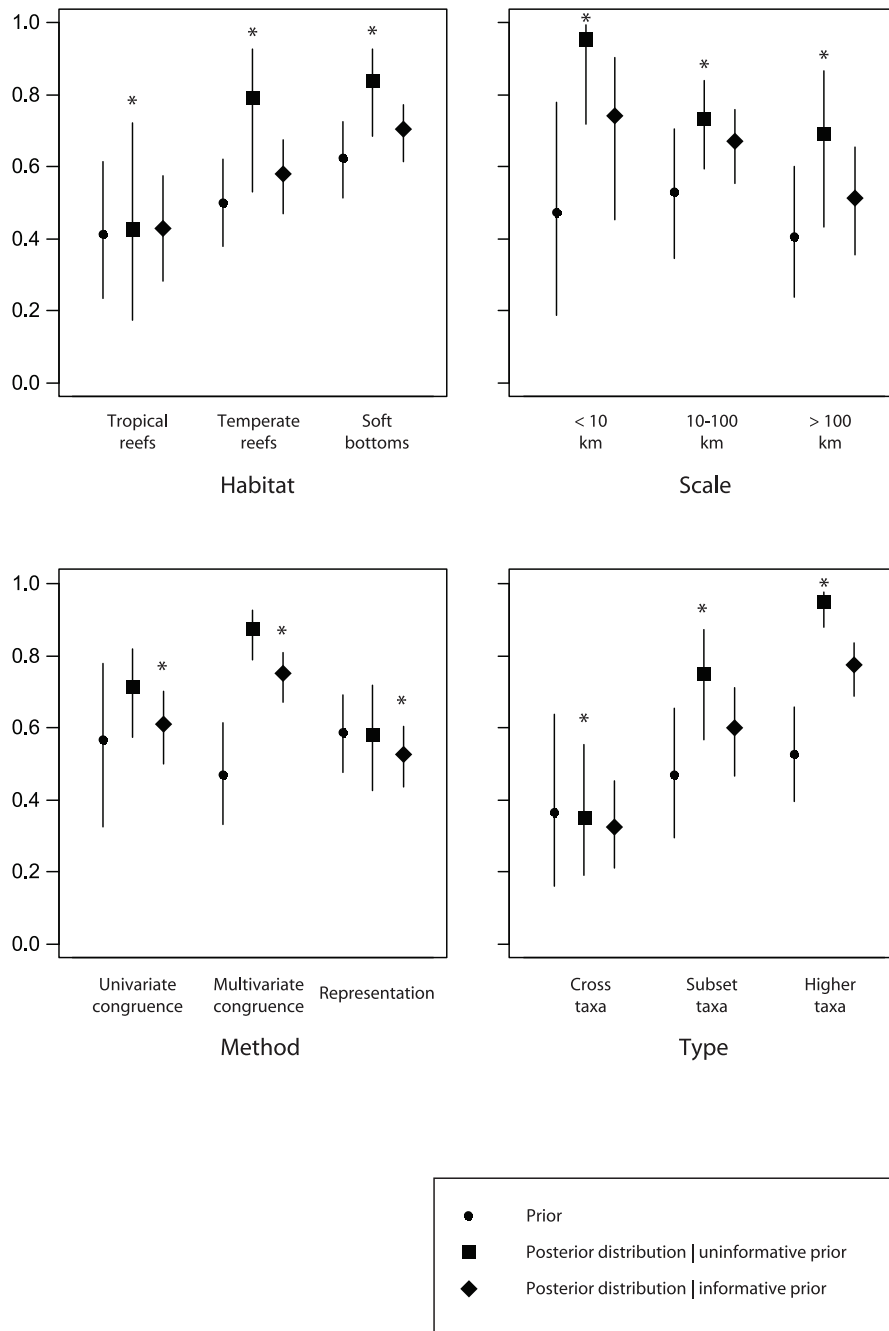


Figure 2. Surrogate effectiveness defined by P , the proportion of tests concluding that surrogate predictions were non-random. Prior distribution (circles), posterior distribution given an uninformative prior (analogous to the likelihood; squares), and posterior distribution given an informative prior (diamonds). Error bars depict the standard deviation of the prior or posterior. Asterisks indicate the best model according to the deviance information criterion. Factors include the marine habitat (*Habitat*), spatial scale (*Scale*), the statistical method used to assess surrogate performance (*Method*) and the type of surrogate (*Type*).
doi:10.1371/journal.pone.0020141.g002

complexity of coral reefs might have been responsible for weaker relationships between surrogates and their target taxa. Indeed, both theoretical and experimental evidence suggest that species

diversity, which is exceptionally high in tropical coral reefs [39,40], is strongly correlated to functional diversity [41,42]. This strong coupling could reflect a high partitioning of resources among

R^2

Parameter estimate

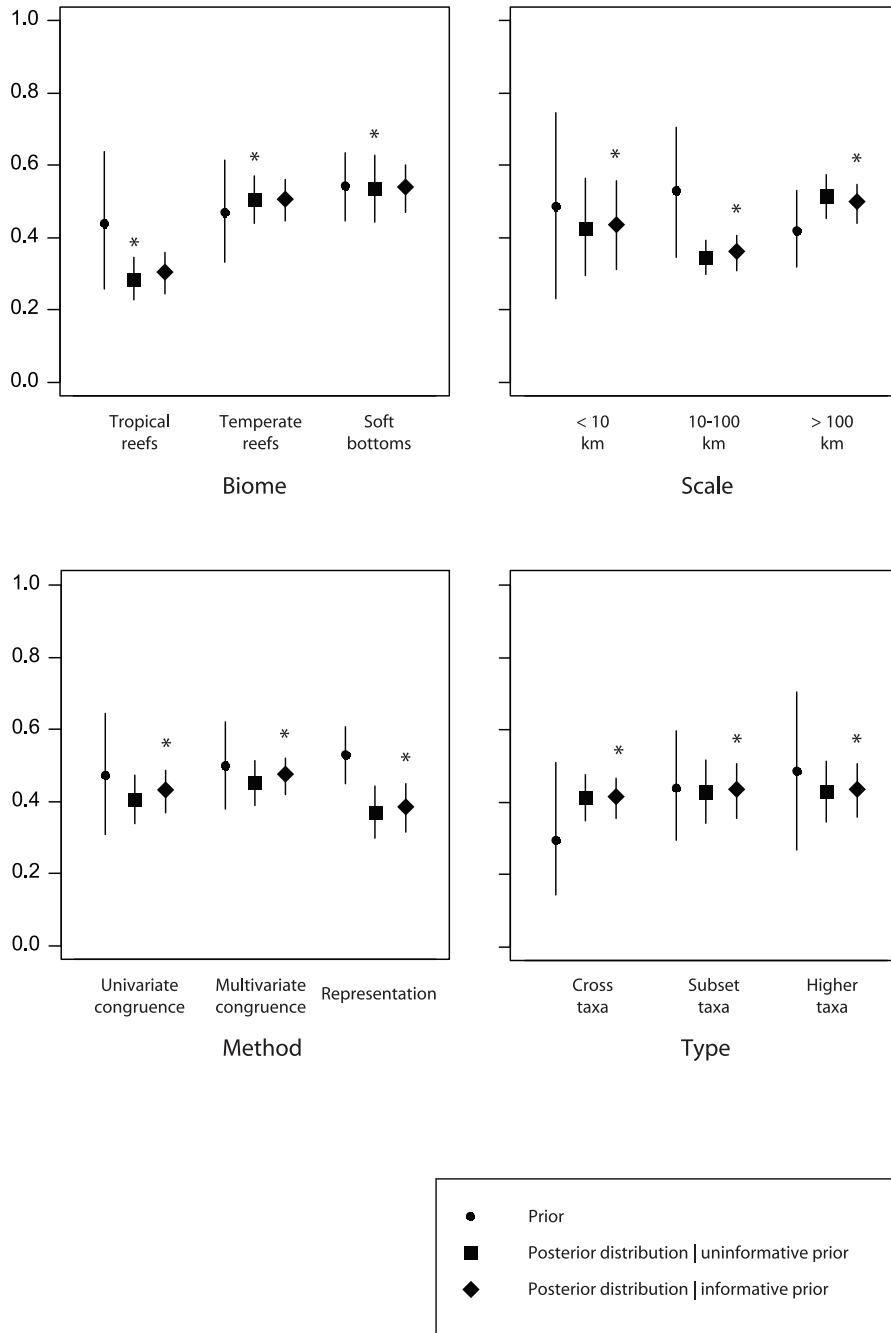


Figure 3. Surrogate effectiveness defined by R^2 , the predictability of targets using surrogates. Prior distribution (circles), posterior distribution given an uninformative prior (analogous to the likelihood; squares) and posterior distribution given an informative prior (diamonds). Error bars depict the standard deviation of the prior or posterior. Asterisks indicate the best model according to the deviance information criterion. Factors include the marine habitat (*Habitat*), spatial scale (*Scale*), the statistical method used to assess surrogate performance (*Method*) and the type of surrogate (*Type*).
doi:10.1371/journal.pone.0020141.g003

component species and low overlap in their functional traits [42]. Thus in tropical reefs, a large number of functional groups often results in more complex food webs. These results corroborate the

idea that high species and functional diversity characterizing tropical coral reefs enhances ecosystem complexity, thus weakening predictive and surrogacy relationships between taxa.

Table 4. Deviance information criterion (DIC) for models of surrogate effectiveness defined as the proportion of tests (P) concluding that surrogate predictions are non-random and as the surrogate predictive power (R^2).

Response	Model	Model type	Priors	DIC	
P	type	mixed		207.31	
	type	mixed	informative	208.96	
	method	mixed	informative	212.66	
	method	mixed		212.76	
	scale	mixed		217.44	
	n	mixed		218.08	
	habitat	mixed		218.19	
	null	mixed		218.26	
	habitat	mixed	informative	218.47	
	scale	mixed	informative	218.59	
	type	fixed		221.51	
	method	fixed		248.06	
	habitat	fixed		253.63	
	scale	fixed		266.84	
	null	fixed		269.33	
	n	fixed		270.99	
	R^2	habitat	mixed		-74.55
		habitat	mixed	informative	-74.11
		n	mixed		-73.14
null		mixed		-71.98	
type		mixed	informative	-69.39	
type		mixed		-69	
method		mixed	informative	-68.06	
method		mixed		-68	
scale		mixed	informative	-66.41	
scale		mixed		-65.83	
scale		fixed		-62.25	
habitat		fixed		-61.61	
n		fixed		-53.35	
method		fixed		-52.98	
null		fixed		-46.28	
type		fixed		-43.81	

Factors include the type of surrogate (type), the statistical method used to assess surrogate performance (method), the marine habitat (habitat), the spatial scale (scale) and the sample size (n). Models are ranked by increasing DIC. doi:10.1371/journal.pone.0020141.t004

Value of Bayesian mixed-effect models in meta-analyses

Using Bayesian hierarchical models in a meta-analysis of surrogate effectiveness allowed us to account for correlation among multiple statistical tests within studies. All hierarchical models ranked higher than the analogous non-hierarchical models, showing the relevance of accounting for such correlation, and suggesting that studies neglecting them should be interpreted with caution (e.g., [17]). Multi-level hierarchical models have been extensively used in socio-economic and medical research, for example, to model the relationship among households of a same city [43], but such models have only recently received interest in ecology (e.g., [44,45]). A frequentist framework could alternatively be used to model such within-group correlations; however, when the number of groups is

small or the multi-level design is complicated, there might not be enough information to estimate variance components precisely, whereas a Bayesian approach would average over the uncertainty in all parameters in the model [43].

Incorporating an informative prior in our models proved valuable in a few cases where estimations of the likelihood based on meta-data only were relatively imprecise, or differences among factor levels were small, such as in the *Method* model of *P*, or in the *Scale*, *Method* or *Type* models of R^2 . In other cases, models using informative priors based on expert opinion and those using an uninformative prior were essentially undistinguishable. This might reflect a relatively diffuse distribution of the prior with respect to that of the estimate based on the meta-data. According to the terminology used by Kuhnert et al (2010), we used a direct-elicitation method, conducted remotely through an online survey. This has the advantage of eliciting opinion from multiple international experts when resource constraints prevent face-to-face interviews [20]. However, the drawback is that no Delphi process could be used, i.e., the process by which mutual feedback among experts promotes the convergence of their opinion to reach a consensus, or where the elicitor can provide immediate feedback to the expert in one-to-one style survey. Whether or not using the Delphi approach, expert opinion can still be prone to biases that can emerge from different sources (see Table 3 in Kuhnert et al. 2010). Among those sources, the linguistic uncertainty (inducing a *cognitive bias*, i.e. misunderstanding of what is required) [21] is probably the most accessible to the elicitor. In our survey, measures were taken to reduce the linguistic uncertainty and included the assistance of a social scientist to design the survey, as well as a series of dry tests and feed-back loops among research group colleagues (Mellin C, unpubl. data.).

Recommendations for the use of biological surrogates

Despite an exhaustive literature search, a careful examination of the metadata revealed that specific surrogate and target groups were not uniformly distributed across spatial scales and habitats. For example, all studies targeting arthropods were done at a spatial scale >100 km (Table S2). We found that the lowest surrogate effectiveness, observed in tropical reefs, was always associated with tests examining corals. Although we could not fit any model that included interactions between taxon and spatial scale (or habitat) because of missing cross-factor combinations, we contend that associations between a specific taxon and a spatial scale (or habitat) did not bias overall results across all taxa considered (total of 16). However, these associations still reflect ecologists' expectations as to the spatial scales or habitats where surrogates should be the most effective, and can in turn be analysed qualitatively to revisit these expectations and inform the sampling design of future studies. Likewise, studies targeting arthropods could possibly benefit from the consideration of a spatial scale <10 km, whereas surrogate effectiveness in coral reefs could be higher when considering taxa other than corals.

We expect surrogate effectiveness to be the greatest for higher-taxon surrogates at a <10-km spatial scale, in low-complexity marine ecosystems such as soft bottoms, and using multivariate-based methods. In addition, surrogate taxa should ideally have a broad distribution across different environments [2] but also incorporate many species with restricted distributions [14,33,46,47], be easy and cost-effective to identify and survey [47], and be amenable to survey at multiple spatial scales [48]. A lack of spatial consistency in surrogacy due, for example, to regional patterns in species distributions, might make indicator groups perform differently among areas [47].

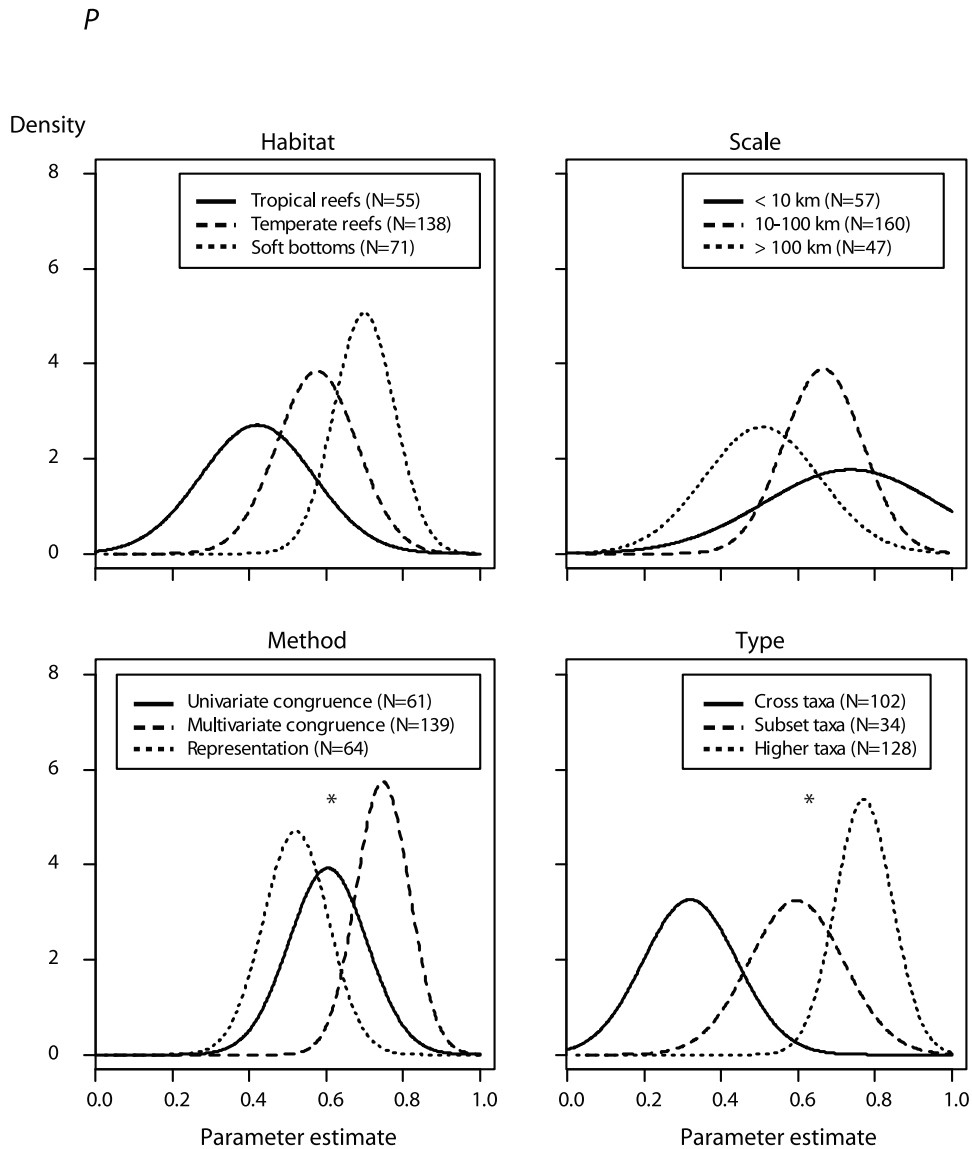


Figure 4. Posterior distributions (given an informative prior) of surrogate effectiveness defined as P . Posterior distributions of P (i.e. the proportion of tests concluding that surrogate predictions are non-random) are given according to the marine habitat (*Habitat*), spatial scale (*Scale*), the statistical method used to assess surrogate performance (*Method*) and the type of surrogate (*Type*). Asterisks indicate models outperforming the null model. A Gaussian distribution with the mean and standard deviation of the posterior distribution was used to approximate posterior distributions.

doi:10.1371/journal.pone.0020141.g004

While often used in combination and compared, the different statistical methods applied to construct the surrogate serve different purposes. One third of the studies examined here used a combination of methods based on spatial congruence and representation to answer the same question, despite the two approaches addressing different issues. While congruence-based methods (uni- or multivariate) are informative for prediction, they cannot inform conservation planning efficiency. Ensuring the cost-effectiveness of conservation efforts requires maximizing biodiversity (or endemism) included in a minimum number of sites [7,49,50]. This can only be achieved using representation-based methods such as reserve-selection algorithms. Therefore, where the goal of a study is the prediction of biodiversity patterns, surrogacy methods based on spatial congruence should be used; where conservation planning is the goal, surrogacy methods based

on representation are applicable. However, such a distinction between the utility of these different approaches to biological surrogates is rarely apparent in the published literature. Future investigation of biological surrogacy will benefit from careful choice and specification of methods depending on whether the goal of the study is prediction or planning.

Independent of the ecosystems in which they are applied, the reliability of biological surrogates can be improved in the following ways: (i) Surrogates should be selected based on the target taxa of interest, with an awareness of the limits imposed by that selection. Higher-taxa surrogates can be appropriate (and often effective) when identification of target taxa to species is expensive, while cross-taxa surrogates are appropriate (but rarely effective) when the target is difficult to census. (ii) The objective of using a particular surrogate needs to be explicit, and the method used for

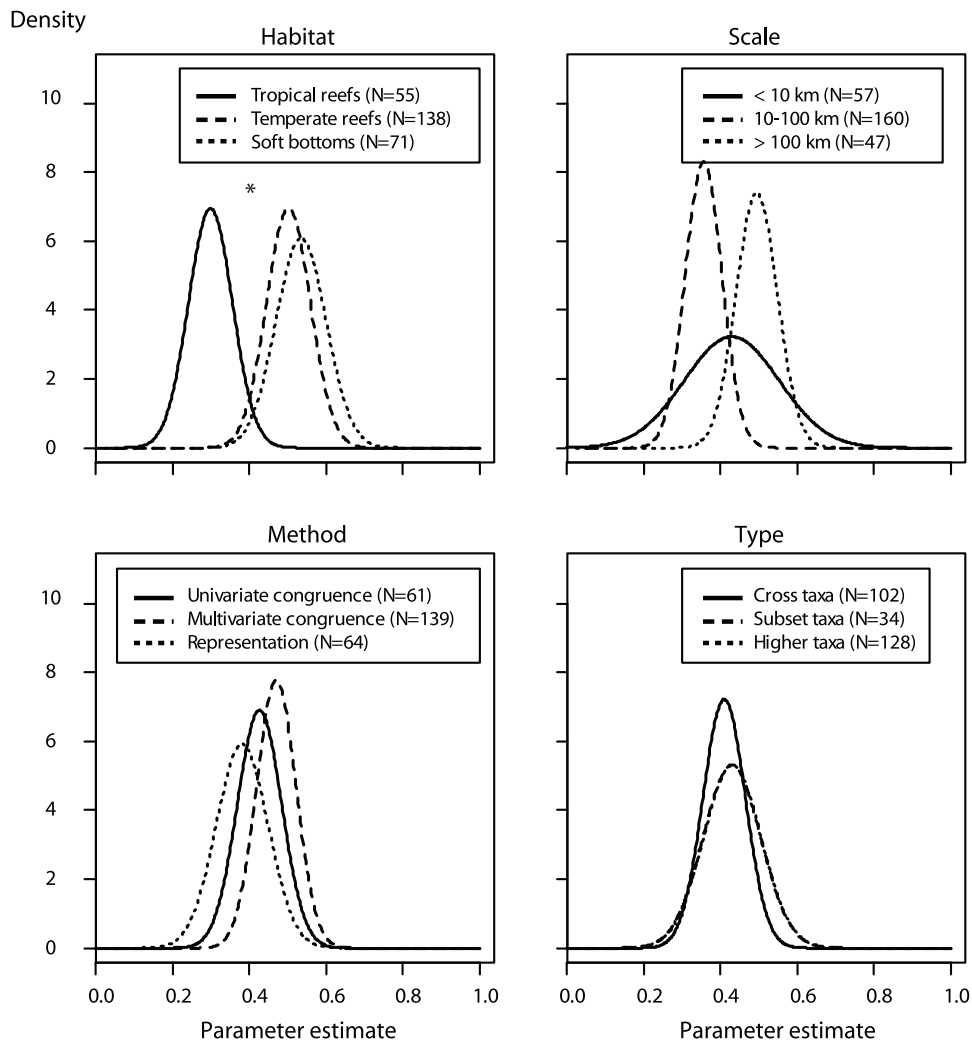
R^2 

Figure 5. Posterior distributions (given an informative prior) of surrogate effectiveness defined as R^2 . Posterior distributions of R^2 (i.e. the surrogate predictive power) are given according to the marine habitat (*Habitat*), spatial scale (*Scale*), the statistical method used to assess surrogate performance (*Method*) and the type of surrogate (*Type*). Asterisks indicate models outperforming the null model. A Gaussian distribution with the mean and standard deviation of the posterior distribution was used to approximate posterior distributions. doi:10.1371/journal.pone.0020141.g005

its construction must be matched to this objective to maximize its effectiveness. If the surrogate is to be used for designing networks of protected areas, only representation-based methods are appropriate, whereas congruence-based methods should be used to predict patterns of biodiversity where data are scarce or unavailable. (iii) The costs of sampling surrogates need to be evaluated to optimize their efficiency. To be efficient, a surrogate must be considerably less expensive and time-consuming than sampling the target [2]. There is, however, currently little information available regarding the costs of monitoring biological surrogates in marine ecosystems (but see [51] for the case of tropical forests). (iv) Finally, the temporal stability of surrogate effectiveness needs to be tested and not just assumed. Most biological surrogacy studies have attempted to evaluate surrogate effectiveness across space only, pooling samples among times, without assessing the temporal robustness of the surrogate-target

relationship (but see [52]). Doing so will be particularly useful in the context of monitoring the species-specific responses to global change. By advancing knowledge in these four areas, a better understanding of the properties of surrogates and how to deploy them most effectively and efficiently will be gained. Bridging these knowledge gaps is crucial as biological surrogates are becoming an increasingly important tool for conservation planning.

Supporting Information

Table S1 Results obtained for the respondents of the online survey available at <http://www.adelaide.edu.au/environment/mbp/survey/02.html>. (DOC)

Table S2 Meta-data and references of literature reviewed. (DOC)

Text S1 WinBUGS code for Bayesian modelling.
(DOC)

Text S2 Literature reviewed.
(DOC)

Acknowledgments

This study is a contribution from the Australian Commonwealth Environment Research Facilities (CERF) programme. Partners in the CERF Marine Biodiversity Hub are the University of Tasmania, CSIRO Wealth from Oceans Flagship, Geoscience Australia, Australian Institute of Marine Science and Museum Victoria. R. P. publishes with permission of

the Chief Executive Officer of Geoscience Australia. We thank A. Fordham for his invaluable help and advice on the survey design. Expert opinion was provided by M. Beger, W. Gladstone, P. Hutchings, L. Kaufman, A. Terlizzi and M. Vanderklift. Thanks to F. Althaus, N. Barrett, D. Roelke and N. Shears for helpful comments.

Author Contributions

Conceived and designed the experiments: CM SD JC CB. Performed the experiments: CM SD. Analyzed the data: CM SD. Wrote the paper: CM SD MJC CB MM. Decided on paper scope and aims: CM JC CB AW CRP GE. Reviewed and contributed to the paper: CRP GE RP AW.

References

- Sarkar S, Margules C (2002) Operationalizing biodiversity for conservation planning. *J Biosci* 27: 299–308.
- Noss RF (1990) Indicators for monitoring biodiversity - A hierarchical approach. *Cons Biol* 4: 355–364.
- Margules CR, Pressey RL, Williams PH (2002) Representing biodiversity: data and procedures for identifying priority areas for conservation. *J Biosci* 27: 309–326.
- Garson J, Aggarwal A, Sarkar S (2002) Birds as surrogates for biodiversity: an analysis of a data set from southern Quebec. *J Biosci* 27: 347–360.
- Rodrigues ASL, Brooks TM (2007) Shortcuts for biodiversity conservation planning: the effectiveness of surrogates. *Annu Rev Ecol Syst* 38: 713–717.
- McArthur MA, Brooke BP, Przeslawski R, Ryan DA, Lucieer VL, et al. (2010) On the use of abiotic surrogates to describe marine benthic biodiversity. *Estuar Coast Shelf Sci* 88: 21–32.
- Lamoreux JF, Morrison JC, Ricketts TH, Olson DM, Dinerstein E, et al. (2006) Global tests of biodiversity concordance and the importance of endemism. *Nature* 440: 212–214.
- Orme CDL, Davies RG, Burgess M, Eigenbrod F, Pickup N, et al. (2005) Global hotspots of species richness are not congruent with endemism or threat. *Nature* 436: 1016–1019.
- Margules CR, Pressey RL (2000) Systematic conservation planning. *Nature* 405: 243–253.
- Hirst AJ (2008) Surrogate measures for assessing cryptic faunal biodiversity on macroalgal-dominated subtidal reefs. *Biol Cons* 141: 211–220.
- Smith SDA (2005) Rapid assessment of invertebrate biodiversity on rocky shores: where there's a whelk there's a way. *Biodivers Conserv* 14: 3565–3576.
- Olgard F, Somerfield PJ (2000) Surrogates in marine benthic investigations: Which taxonomic unit to target? *J Aquat Ecosyst Stress Recovery* 7: 25–42.
- Ball IR, Possingham HP, Watts M (2009) Marxan and relatives: Software for spatial conservation prioritisation. In: Moilanen A, Wilson KA, Possingham HP, eds. *Spatial conservation prioritisation: Quantitative methods and computational tools*. Oxford, UK: Oxford University Press.
- Beger M, Jones GP, Munday PL (2003) Conservation of coral reef biodiversity: a comparison of reserve selection procedures for corals and fishes. *Biol Cons* 111: 53–62.
- Ward TJ, Vanderklift MA, Nicholls AO, Kenchington RA (1999) Selecting marine reserves using habitats and species assemblages as surrogates for biological diversity. *Ecol Appl* 9: 691–698.
- Poore GCB, Wilson GDF (1993) Marine species richness. *Nature* 361: 597–598.
- Lewandowski AS, Noss RF, Parson DR (2010) The effectiveness of surrogate taxa for the representation of biodiversity. *Cons Biol* 24: 1367–1377.
- McCarthy MA (2007) *Bayesian methods for ecology*. Cambridge: University Press.
- Moller AP, Jennions MD (2001) Testing and adjusting for publication bias. *Trends Ecol Evol* 16: 580–586.
- Kuhnert PM, Martin TG, Griffiths SP (2010) A guide to eliciting and using expert knowledge in Bayesian ecological models. *Ecol Lett* 13: 900–914.
- Low-Choy S, O'Leary R, Mengersen K (2009) Elicitation by design in ecology: using expert opinion to inform priors for Bayesian statistical models. *Ecology* 90: 265–277.
- Burnham KP, Anderson DR (2002) *Model selection and multimodel inference: a practical information theoretic approach*, 2nd ed. New York: Springer-Verlag.
- Egger M, Smith GD, Schneider M, Minder C (1997) Bias in meta-analysis detected by a simple, graphical test. *Br Med J* 315: 629–634.
- Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. *Austral Ecol* 26: 32–46.
- Branscum AJ, Johnson WO, Thurmond MC (2007) Bayesian beta regression: applications to household expenditure data and genetic distance between foot-and-mouth disease viruses. *Australian and New Zealand Journal of Statistics* 49: 287–301.
- Spiegelhalter D, Best N, Carlin B, van der Linde A (2002) Bayesian measures of model complexity and fit. *J Roy Stat Soc B* 64: 583–639.
- Spiegelhalter D, Thomas A, Best NG, Lunn D (2003) WinBUGS Version 1.4 Imperial College & MRC Biostatistics Unit.
- Sturtz S, Ligges U, Gelman A (2005) R2WinBUGS: A package for running WinBUGS from R. *Journal of Statistical Software* 12: 1–13.
- R Development Core Team (2009) R: A language and environment for statistical computing. Vienna, Austria. URL <http://www.R-project.org/>: R Foundation for Statistical Computing.
- Williams PH, Gaston KJ (1994) Measuring more of biodiversity - Can higher-taxon richness predict wholesale species richness. *Biol Cons* 67: 211–217.
- Anderson MJ, Diebel CE, Blom WM, Landers TJ (2005) Consistency and variation in kelp holdfast assemblages: Spatial patterns of biodiversity for the major phyla at different taxonomic resolutions. *J Exp Mar Biol Ecol* 320: 35–56.
- Bertrand Y, Ptejfel F, Rouse GW (2006) Taxonomic surrogacy in biodiversity assessments, and the meaning of Linnaean ranks. *Systematics and Biodiversity* 4: 149–159.
- Williams P, Faith D, Manne L, Sechrest W, Preston C (2006) Complementarity analysis: Mapping the performance of surrogates for biodiversity. *Biol Cons* 128: 253–264.
- Ferrier S (2002) Mapping spatial pattern in biodiversity for regional conservation planning: Where to from here? *Syst Biol* 51: 331–363.
- Gotelli NL, Colwell RK (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecol Lett* 4: 379–391.
- Howard PC, Viskanic P, Davenport TRB, Kigenyi FW, Baltzer M, et al. (1998) Complementarity and the use of indicator groups for reserve selection in Uganda. *Nature* 394: 472–475.
- Moritz C, Richardson KS, Ferrier S, Monteith GB, Stanicic J, et al. (2001) Biogeographical concordance and efficiency of taxon indicators for establishing conservation priority in a tropical rainforest biota. *Proc R Soc Lond B* 268: 1875–1881.
- Gaston KJ (2000) Global patterns in biodiversity. *Nature* 405: 220–227.
- Connell JH (1978) Diversity in tropical rain forest and coral reef. *Science* 199: 1302–1310.
- Knowlton N, Brainard RE, Fisher R, Moews M, Plaisance L, et al. (2010) Coral reef biodiversity. In: McIntyre A, ed. *Life in the World's Oceans: Diversity, Distribution, and Abundance*. Chichester: Wiley-Blackwell.
- Nacem S (2002) Disentangling the impacts of diversity on ecosystem functioning in combinatorial experiments. *Ecology* 83: 2925–2935.
- Micheli F, Halpern BS (2005) Low functional redundancy in coastal marine assemblages. *Ecol Lett* 8: 391–400.
- Gelman A, Hill J (2007) *Data analysis using regression and multilevel hierarchical models*. New York: Cambridge University Press.
- MacNeil MA, Graham NAJ, Polunin NVC, Kulbicki M, Galzin R, et al. (2009) Hierarchical drivers of reef-fish metacommunity structure. *Ecology* 90: 252–264.
- Qian SS, Cuffney TF, Alameddine I, McMahon G, Reckhow KH (2010) On the application of multilevel modeling in environmental and ecological studies. *Ecology* 91: 355–361.
- Caro TM, O'Doherty G (1999) On the use of surrogate species in conservation biology. *Cons Biol* 13: 805–814.
- Lawler JJ, White D, Sifneos JC, Master LL (2003) Rare species and the use of indicator groups for conservation planning. *Cons Biol* 17: 875–882.
- Sauberer N, Zulka KP, bentsperg-Traun M, Berg HM, Bieringer G, et al. (2004) Surrogate taxa for biodiversity in agricultural landscapes of eastern Austria. *Biol Cons* 117: 181–190.
- Brooks T, Balmford A, Burgess N, Fjeldsa J, Hansen LA, et al. (2001) Toward a blueprint for conservation in Africa. *BioScience* 51: 613–624.
- Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.
- Gardner TA, Barlow J, Araujo IS, vila-Pires TC, Bonaldo AB, et al. (2008) The cost-effectiveness of biodiversity surveys in tropical forests. *Ecol Lett* 11: 139–150.
- Magierowski RH, Johnson CR (2006) Robustness of surrogates of biodiversity in marine benthic communities. *Ecol Appl* 16: 2264–2275.