

Universidade Federal de Goiás
Instituto de Ciências Biológicas
Programa de Pós-Graduação em Ecologia e Evolução

Oportunidades e desafios em macroecologia marinha

ANDRÉ MENEGOTTO DOMINGOS

GOIÂNIA

2019

**TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS
DE TESES E
DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG**

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

1. Identificação do material bibliográfico: ☐ Dissertação ☒ Tese

2. Identificação da Tese ou Dissertação:

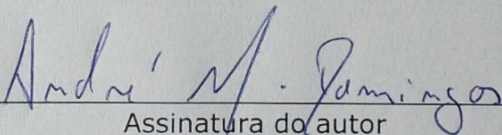
Nome completo do autor: André Menegotto Domingos

Título do trabalho: Oportunidades e desafios em macroecologia marinha

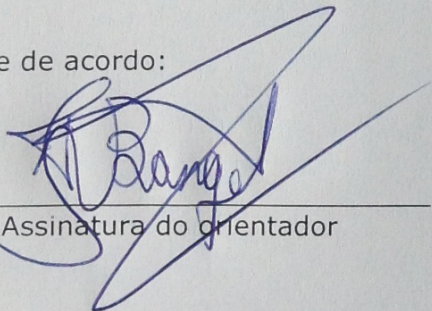
3. Informações de acesso ao documento:

Concorda com a liberação total do documento ☒ SIM ☐ NÃO

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF da tese ou dissertação.


Assinatura do autor

Ciente e de acordo:


Assinatura do orientador

Data: 27 / 03 / 2019

ANDRÉ MENEGOTTO DOMINGOS

Oportunidades e desafios em macroecologia marinha

Tese apresentada ao Programa de Pós-Graduação em Ecologia e Evolução do Departamento de Ecologia do Instituto de Ciências Biológicas da Universidade Federal de Goiás como requisito parcial para a obtenção do título de Doutor em Ecologia e Evolução.

Orientador: Prof. Dr. Thiago Fernando Rangel

GOIÂNIA

2019

Ficha de identificação da obra elaborada pelo autor, através do
Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Domingos, André Menegotto
Oportunidades e desafios em macroecologia marinha [manuscrito]
/ André Menegotto Domingos. - 2019.
182 f.

Orientador: Prof. Dr. Thiago Fernando Lopes Valle de Britto
Rangel.
Tese (Doutorado) - Universidade Federal de Goiás, Instituto de
Ciências Biológicas (ICB), Programa de Pós-Graduação em Ecologia e
Evolução, Goiânia, 2019.
Bibliografia. Apêndice.
Inclui siglas, mapas, gráfico, tabelas.

1. Biodiversidade marinha. 2. Gradiente latitudinal de diversidade.
3. Produtividade. 4. Riqueza de espécies. 5. Viés amostral. I. Rangel,
Thiago Fernando Lopes Valle de Britto, orient. II. Título.

CDU 574



**SERVIÇO PÚBLICO FEDERAL
UNIVERSIDADE FEDERAL DE GOIÁS - UFG
INSTITUTO DE CIÊNCIAS BIOLÓGICAS - ICB
PROGRAMA DE PÓS-GRADUAÇÃO EM ECOLOGIA E EVOLUÇÃO - PPGEE**

ATA DA SESSÃO PÚBLICA DE DEFESA DE TESE Nº 78

Aos vinte e sete dias do mês de fevereiro de 2019 (27/02/2019), às oito horas (08h), no Auditório do ICB V, UFG, reuniram-se os componentes da banca examinadora: **Prof. Dr. Thiago Fernando Lopes Valle de Britto Rangel, ICB/UFG; Prof. Dr. Matheus de Souza Lima Ribeiro, ICB/UFG; Prof. Dr. José Alexandre Felizola Diniz Filho, ICB/UFG; Dra. Geiziane Tassarolo, UEG; Prof. Dr. Sérgio Ricardo Floeter, UFSC;** para, em sessão pública presidida pelo (a) primeiro(a) examinador(a) citado(a), procederem à avaliação da defesa de tese intitulada: **"Oportunidades e desafios em macroecologia marinha"**, em nível de doutorado, área de concentração em Ecologia e Evolução, de autoria de **André Menegotto Domingos**, discente do Programa de Pós-Graduação em Ecologia e Evolução da Universidade Federal de Goiás. A sessão foi aberta pelo(a) presidente(a), que fez a apresentação formal dos membros da banca. A palavra, a seguir, foi concedida a(o) autor(a) da tese que, em cerca de 35 minutos, procedeu à apresentação de seu trabalho. Terminada a apresentação, cada membro da banca arguiu a(o) examinada(o), tendo-se adotado o sistema de diálogo sequencial. Terminada a fase de arguição, procedeu-se à avaliação da tese. Tendo-se em vista o que consta na Resolução nº 1127 de dezembro de 2012 do Conselho de Ensino, Pesquisa, Extensão e Cultura (CEPEC), que regulamenta o Programa de Pós-Graduação em Ecologia e Evolução, a tese foi APROVADA, considerando-se integralmente cumprido este requisito para fins de obtenção do título de Doutor(a) em Ecologia e Evolução pela Universidade Federal de Goiás. A conclusão do curso dar-se-á quando da entrega da versão definitiva da tese na secretaria do programa, com as devidas correções sugeridas pela banca examinadora, no prazo de trinta dias a contar da data da defesa. Cumpridas as formalidades de pauta, às 11 h e

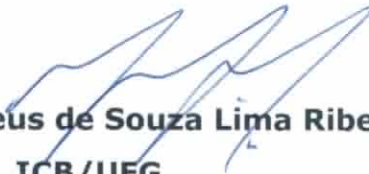
56 min., encerrou-se a sessão de defesa e, para constar, eu, Suely Ana Ribeiro, secretária executiva da Universidade Federal de Goiás - UFG, lavrei a presente ata que, após lida e aprovada, será assinada pelos membros da banca examinadora em três vias de igual teor.



Prof. Dr. Thiago Fernando Lopes Valle de Britto Rangel

Presidente da banca

ICB/UFG



Prof. Dr. Matheus de Souza Lima Ribeiro

ICB/UFG



Prof. Dr. José Alexandre Felizola Diniz Filho

ICB/UFG



Dra. Geiziane Tessarolo

UEG



Prof. Dr. Sérgio Ricardo Floeter

UFSC

“– Data! Data! Data! I can't make bricks without clay.”

Sherlock Holmes em The Adventure of the Copper Beeches

por Sir Arthur Conan Doyle

AGRADECIMENTOS

Ao meu orientador Dr. Thiago Rangel, pela confiança, liberdade e imenso suporte. Muito obrigado, Thiago. Você é um exemplo para mim.

Ao Programa de Pós-Graduação em Ecologia e Evolução, por fornecer um ambiente cientificamente rico, estimulante e agradável.

À CAPES, pela bolsa de estudos sem a qual seria impossível concluir o doutorado.

Aos pesquisadores Julian Schrader, Patrick Weigelt, Holger Kreft, Derek Tittensor e Robert K. Colwell pela colaboração em alguns dos capítulos.

Aos professores Dr. José Alexandre Felizola Diniz Filho, Dr. Matheus de Souza Lima Ribeiro, Dra. Geiziane Tessarolo, Dr. Sérgio Ricardo Floeter, por aceitarem o convite de participar da banca avaliadora.

À minha mãe, Marlene, e minha esposa, Márcia, por todo carinho e apoio.

SUMÁRIO

RESUMO.....	7
ABSTRACT.....	8
INTRODUÇÃO GERAL	9
Referências	13
CAPÍTULO 1: <i>A global test of the subsidized island biogeographic hypothesis</i>	16
Abstract	17
Introduction.....	18
Methods.....	21
Results	26
Discussion	28
References.....	33
CAPÍTULO 2: <i>Taking a marine perspective on the latitudinal diversity gradient</i>	38
Abstract	39
A challenging and fascinating mystery	40
A brief history of the latitudinal diversity gradient in the sea	42
Hypothetical explanations under a marine perspective	43
Concluding remarks	52
References.....	58
CAPÍTULO 3: <i>Mapping knowledge gaps in marine diversity reveals a latitudinal gradient of missing species richness</i>	67
Abstract	68
Introduction.....	69
Results	71
Discussion	75
Methods.....	79
References.....	82
CAPÍTULO 4: <i>Sampling simulation in a virtual ocean reveals strong sample effect in marine diversity patterns</i>	86
Abstract	87
Introduction.....	88
Methods.....	91
Results	95
Discussion	102
References.....	109
CONSIDERAÇÕES FINAIS.....	113
APÊNDICES.....	114
Capítulo 1	115
Capítulo 3.....	138
Capítulo 4.....	174

Resumo

A macroecologia, assim como outros campos da ecologia, tem sido historicamente uma disciplina terrestre. No entanto, a biodiversidade marinha pode conter características fundamentais para a nossa compreensão dos padrões globais de diversidade. Aqui, eu explorei essa possibilidade integrando conhecimento sobre o ambiente marinho com hipóteses e modelos verbais utilizados para explicar variações geográficas na diversidade de espécies. Inicialmente, eu utilizei dados de produtividade marinha para avaliar o efeito de subsídios espaciais na diversidade de plantas insulares (capítulo 1). Esse estudo mostrou que, ao contrário da hipótese proposta, não existem evidências de que recursos alóctones influenciam a diversidade de ilhas pequenas. Posteriormente, eu realizei uma revisão conceitual dos principais mecanismos propostos para explicar o gradiente latitudinal de especiação sob uma perspectiva marinha (capítulo 2). Nessa revisão eu mostro que, ao contrário do sistema terrestre, cada mecanismo gera uma predição totalmente diferente para os oceanos. No entanto, é preciso também reconhecer e destacar as limitações que são inerentes aos dados de biodiversidade marinha, principalmente o viés amostral. Por isso, eu avaliei aqui nosso grau de conhecimento sobre a distribuição latitudinal de diferentes grupos taxonômicos (capítulo 3). Nesse estudo eu revelo a existência de um gradiente de ausência de espécies que é fortemente associado com baixo esforço amostral nos trópicos. Por fim, eu testei através de um modelo de simulação a eficiência do esforço amostral empírico em detectar acuradamente padrões de diversidade gerados em um oceano virtual (capítulo 4). O resultado mostrou que o gradiente de diversidade atualmente observado nos oceanos não difere do que é registrado ao aplicar a mesma amostragem em um simples modelo nulo. Com isso, eu concluo que a biodiversidade marinha apresenta, de fato, um potencial enorme para a compreensão de fenômenos macroecológicos. No entanto, existem ainda algumas lacunas geográficas e taxonômicas que precisam ser preenchidas para que esse conhecimento possa ser explorado em todo seu potencial.

Palavras-chave: Biodiversidade marinha, especiação, gradiente latitudinal de diversidade, oceano, produtividade, riqueza de espécies, viés amostral.

Abstract

Macroecology, like other fields of ecology, has been historically a terrestrial discipline. However, marine biodiversity may contain key features for our understanding of global diversity patterns. Here, I explored this possibility by integrating knowledge about the marine environment with hypotheses and verbal models used to explain geographic variations in species diversity. Initially, I used marine productivity data to evaluate the effect of spatial subsidies on island plant diversity (Chapter 1). This study showed that, contrary to the proposed hypothesis, there is no evidence that allochthonous resources influence the diversity of small islands. Subsequently, I realized a conceptual review of the main mechanisms proposed to explain the latitudinal gradient of speciation under a marine perspective (chapter 2). In this review I show that, unlike the terrestrial system, each mechanism generates predictions totally different in the oceans. Nevertheless, it is also necessary to recognize and highlight the limitations that are inherent to marine biodiversity data, especially sample bias. Therefore, I evaluated here the state of our knowledge about the latitudinal distribution of different taxonomic groups (Chapter 3). In this study I reveal the existence of a species absence gradient that is strongly associated with low sampling effort in the tropics. Finally, I tested through a simulation model the efficiency of the empirical sample effort to accurately detect diversity patterns generated in a virtual ocean (Chapter 4). The result showed that the diversity gradient currently observed in the oceans does not differ from what is recorded after applying the same sampling effort in a simple null model. Therefore, I conclude that marine biodiversity may, indeed, be of great importance to understand macroecological patterns. However, there are still some geographic and taxonomic gaps that need to be addressed in order to explore this knowledge to its full potential.

Keywords: Latitudinal diversity gradient, marine biodiversity, ocean, productivity, sampling bias, speciation, species richness.

Introdução geral

Macroecologia é uma disciplina relativamente recente, tendo sido formalmente descrita e estabelecida apenas no final da década de 80 (Brown & Maurer 1989). Obviamente, isso não significa que estudos em macroecologia não existiram previamente. Padrões macroecológicos já eram descritos pelos primeiros naturalistas no século XIX, incluindo suas possíveis explicações mecanísticas (Brown 1995). No entanto, é inegável o rápido avanço que vem ocorrendo nessa área nos últimos anos, principalmente com o aumento na disponibilidade de bancos de dados sobre a distribuição das espécies, filogenias moleculares, e o crescente poder computacional (Keith *et al.* 2012). Além do aumento na disponibilidade e resolução dos dados, as pesquisas em macroecologia têm também se destacado por cada vez mais se desvencilhar de suas raízes descritivas e correlativas para se tornar uma ciência caracterizada por modelos conceituais baseados em processos (Keith *et al.* 2012). Exemplos incluem tanto o uso de modelos para dissociar padrões que emergem de processos ecológicos das expectativas nulas (Colwell & Hurtt 1994; Jetz & Rahbek 2001), quanto o uso de simulações computacionais envolvendo mecanismos ecológicos e evolutivos que buscam desvendar as variáveis e os parâmetros que melhor explicam os padrões observados na natureza (Rangel *et al.* 2007, 2018; Coelho *et al.* 2018).

Apesar do escopo global e dos recentes avanços, umas das barreiras ainda a serem ultrapassadas pela macroecologia é que ela tem sido historicamente uma disciplina alimentada por dados e conceitos gerados principalmente no ambiente terrestre (Witman & Roy 2009; Beck *et al.* 2012). Uma rápida busca nos artigos do periódico *Global Ecology and Biogeography*, por exemplo, mostrou que dentre os seus 1.600 artigos já publicados apenas 88 possuem os termos “marine OR ocean* OR *sea OR coast* OR reef” em seus títulos (busca realizada em 29/01/2019). Certamente essa é uma medida subestimada, já que nem

todos os artigos devem trazer termos relacionados ao bioma nos seus títulos. No entanto, o dobro de artigos (191) foi encontrado utilizando apenas o termo “forest*”. Portanto, é bem provável que uma pesquisa mais refinada retorne uma proporção muito parecida, ou ainda mais desigual.

Mesmo com esta representação numericamente modesta, estudos com organismos marinhos têm feito uma contribuição substancial na área de macroecologia. Alguns avanços na compreensão de padrões globais de diversidade, por exemplo, têm sido realizados com uso de dados sobre invertebrados fósseis (Jablonski *et al.* 2006; Yasuhara *et al.* 2017), interações ecológicas (Longo *et al.* 2019) e dinâmica ambiental em modelos de simulação (Brayard *et al.* 2005; Tittensor & Worm 2016). Esse teste de hipóteses em ambos os sistemas é essencial para o avanço da macroecologia, pois garante uma constante reavaliação das teorias ecológicas enquanto avança nossa compreensão sobre as especificidades de cada sistema (Munguia & Ojanguren 2015). Um ótimo exemplo disso é o recente estudo de Grady *et al.* (2019), que mostra como a energia ambiental influencia a interação entre espécies endotérmicas e ectotérmicas e, por sua vez, determina a diversidade global de predadores nos oceanos. No entanto, existe ainda uma série de padrões biológicos e ambientais que podem ser utilizados para ampliar nossa compreensão dos processos ecológicos, mas que continuam sub-explorados. As contrastantes características de história de vida de muitos dos taxa marinhos, por exemplo, oferecem oportunidades únicas de testar processos causais (Webb 2012). Nesse sentido, os padrões de biodiversidade marinha podem ser interpretados como uma fonte complementar de dados ecológicos, enriquecendo cada vez mais nosso conhecimento sobre a biodiversidade global.

Eu inicio a tese mostrando justamente dois exemplos de como o conhecimento sobre padrões marinhos pode ser útil para testar e avaliar efetivamente hipóteses sobre a variação espacial no número de espécies. No Capítulo 1, eu integro dados de riqueza de angiospermas

em ilhas oceânicas com dados de produtividade marinha para avaliar a hipótese da biogeografia de ilhas subsidiadas (Anderson & Wait 2001). Essa hipótese foi proposta como uma alternativa para explicar o desvio da relação espécie-área em ilhas pequenas (ilhas *stricto sensu* ou habitats ilha), mas permaneceu praticamente sem ser testada desde sua publicação. Aqui, eu testei se a produtividade adjacente as ilhas poderia funcionar como um subsídio para as angiospermas insulares e influenciar seu número esperado de espécies.

No entanto, um dos grandes objetivos da macroecologia ainda é compreender a origem do gradiente latitudinal de diversidade. Embora uma série de hipóteses tenham sido propostas para explicar esse padrão ao longo das últimas décadas (Pianka 1966; Mittelbach *et al.* 2007), existe pouco consenso sobre quais mecanismos efetivamente geram esse padrão e quais são apenas agentes secundários (Hawkins 2004; mas veja Hawkins 2008). Por isso, no Capítulo 2 eu apresento uma revisão conceitual sobre as principais hipóteses evolutivas que foram propostas para explicar o gradiente de diversidade, recorrendo e contrastando seus pressupostos com dados empíricos do ambiente marinho. O interessante aqui, como você perceberá ao longo da leitura, é que muitos padrões físicos e biológicos que são normalmente espacialmente correlacionados com a latitude no ambiente terrestre possuem uma distribuição reversa ou não linear no ambiente marinho. Assim, assumindo que o gradiente de diversidade seja similar entre os dois sistemas (Hillebrand 2004; Kinlock *et al.* 2018), algumas hipóteses inevitavelmente terão maior poder de generalização que outras ao cruzar a interface terra-mar.

A grande desvantagem sobre o uso de dados marinhos para compreender padrões de diversidade é que como o oceano é extremamente vasto, cobrindo 70,8% de toda a superfície terrestre, grande parte de sua biodiversidade permanece inacessível ou com uma acessibilidade muito restrita. Isso faz com que muito de nosso conhecimento sobre a vida marinha seja ainda incompleto (Boeuf 2011), o que pode levar a falsas interpretações da

realidade. Um ótimo exemplo disso é o famoso caso da hipótese azóica, proposta por Edward Forbes em 1844. Após coletar amostras do Mar Mediterrâneo a bordo no navio HMS *Beacon*, Forbes propôs que o fundo marinho era desprovido de vida abaixo dos 550 m, algo que fez muito sentido para sua época (Tyler 2003a). O que ele e outros pesquisadores não sabiam é que essa era uma exceção, não a regra. Condições muito particulares como falta de espécies colonizadoras, temperatura relativamente alta para águas profundas (~13 °C) e baixo fluxo vertical de matéria orgânica contribuem para baixa biodiversidade no fundo do Mar Mediterrâneo (Tyler 2003b). No entanto, essa hipótese perdurou por cerca de 30 anos. Apenas a contínua pesquisa revelou que o mar profundo possui, na verdade, uma fauna rica e variada. Nesse sentido, se não formos cautelosos e identificarmos nossa limitação de informações sobre os oceanos corremos o risco de cair em armadilhas conceituais, criando hipóteses e debatendo possíveis explicações sobre um falso padrão. Por isso, eu dedico os dois capítulos finais da tese a um dos principais fantasmas da macroecologia marinha: o viés amostral.

A distribuição do esforço amostral é extremamente desigual ao longo dos oceanos. De modo geral, algumas poucas regiões de altas latitudes possuem uma cobertura amostral muito mais ampla e intensa que regiões tropicais (Mora *et al.* 2008; Fernandez & Marques 2017). Mesmo assim, os efeitos concretos desse viés sobre os padrões globais de diversidade marinha continuam sendo pouco conhecidos. No Capítulo 3, eu exploro mais a fundo esses efeitos ao analisar a ausência das espécies ao longo das latitudes. Mais especificamente, eu quantifiquei os gaps latitudinais que existem na distribuição geográfica de diversas espécies marinhas, variando de pequenos copepodes até grandes tubarões. O resultado, como você verá, é um gradiente latitudinal de espécies ausentes com pico no equador.

Embora a redução de esforço amostral esteja altamente associada com a ausência de espécies, ainda não é possível afirmar o quanto esse viés está, de fato, influenciando nossa

percepção sobre a variação de diversidade ao longo das latitudes. Para avaliar o efeito do viés amostral na distribuição das espécies e nos padrões de diversidade, no Capítulo 4 eu desenvolvi um modelo de simulação que integra modelos nulos em macroecologia com a distribuição empírica dos eventos amostrais para replicar o processo de amostragem em um mundo virtual previamente conhecido. Com isso, eu pude avaliar o quanto os padrões descritos com base nos dados observados são suficientemente robustos ao viés amostral ou, ao contrário, se eles representam apenas a distribuição espacial da nossa ignorância sobre a diversidade marinha.

Referências

- Anderson, W.B. & Wait, D.A. (2001). Subsidized island biogeography hypothesis: another new twist on an old theory. *Ecol. Lett.*, 4, 289–291.
- Beck, J., Ballesteros-Mejia, L., Buchmann, C.M., Dengler, J., Fritz, S.A., Gruber, B., *et al.* (2012). What's on the horizon for macroecology? *Ecography*, 35, 673–683.
- Boeuf, G. (2011). Marine biodiversity characteristics. *C. R. Biol.*, 334, 435–440.
- Brayard, A., Escarguel, G. & Bucher, H. (2005). Latitudinal gradient of taxonomic richness: combined outcome of temperature and geographic mid-domains effects? *J. Zool. Syst. Evol. Res.*, 43, 178–188.
- Brown, J.H. (1995). *Macroecology*. University of Chicago Press, Chicago.
- Brown, J.H. & Maurer, B.A. (1989). Macroecology: the division of food and space among species on continents. *Science*, 243, 1145–1150.
- Coelho, M.T.P., Dambros, C., Rosauer, D.F., Pereira, E.B. & Rangel, T.F. (2018). Effects of neutrality and productivity on mammal richness and evolutionary history in Australia. *Ecography*, in press.
- Colwell, R.K. & Hurtt, G.C. (1994). Nonbiological gradients in species richness and a spurious Rapoport effect. *Am. Nat.*, 144, 570–595.
- Fernandez, M.O. & Marques, A.C. (2017). Diversity of diversities: a response to Chaudhary, Saeedi, and Costello. *Trends Ecol. Evol.*, 32, 232–234.

- Grady, J. M., Maitner, B.S., Winter, A.S., Kaschner, K., Tittensor, D.P., Record, S., *et al.* (2019). Metabolic asymmetry and the global diversity of marine predators. *Science*, 363, eaat4220.
- Hawkins, B.A. (2004). Are we making progress toward understanding the global diversity gradient? *Basic Appl. Ecol.*, 5, 1–3.
- Hawkins, B.A. (2008). Perspectives in biogeography: recent progress toward understanding the global diversity gradient. *IBS Newsl.*, 6, 5–8.
- Hillebrand, H. (2004). On the generality of the latitudinal diversity gradient. *Am. Nat.*, 163, 192–211.
- Jablonski, D., Kaustuv, R. & Valentine, J.W. (2006). Out of the tropics: evolutionary diversity gradient. *Science.*, 314, 102–106.
- Jetz, W. & Rahbek, C. (2001). Geometric constraints explain much of the species richness pattern in African birds. *Proc. Natl. Acad. Sci.*, 98, 5661–5666.
- Keith, S.A., Webb, T.J., Böhning-Gaese, K., Connolly, S.R., Dulvy, N.K., Eigenbrod, F., *et al.* (2012). What is macroecology? *Biol. Lett.*, 8, 904–906.
- Kinlock, N.L., Prowant, L., Herstoff, E.M., Foley, C.M., Akin-Fajiye, M., Bender, N., *et al.* (2018). Explaining global variation in the latitudinal diversity gradient: meta-analysis confirms known patterns and uncovers new ones. *Glob. Ecol. Biogeogr.*, 27, 125–141.
- Longo, G.O., Hay, M.E., Ferreira, C.E.L. & Floeter, S.R. (2019). Trophic interactions across 61 degrees of latitude in the Western Atlantic. *Glob. Ecol. Biogeogr.*, 28, 107–117.
- Mittelbach, G.G., Schemske, D.W., Cornell, H. V., Allen, A.P., Brown, J.M., Bush, M.B., *et al.* (2007). Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecol. Lett.*, 10, 315–331.
- Mora, C., Tittensor, D.P. & Myers, R.A. (2008). The completeness of taxonomic inventories for describing the global diversity and distribution of marine fishes. *Proc. R. Soc. B Biol. Sci.*, 275, 149–155.
- Munguia, P. & Ojanguren, A.F. (2015). Bridging the gap in marine and terrestrial studies. *Ecosphere*, 6, 1–4.
- Pianka, E.R.. (1966). Latitudinal gradients in species diversity: a review of concepts. *Am. Nat.*, 100, 33–46.
- Rangel, T.F., Edwards, N.R., Holden, P.B., Diniz-Filho, J.A.F., Gosling, W.D., Coelho, M.T.P., *et al.* (2018). Modeling the ecology and evolution of biodiversity: biogeographical cradles, museums, and graves. *Science*, 361, eaar5452.

- Rangel, T.F.L.V.B., Diniz-Filho, J.A.F. & Colwell, R.K. (2007). Species richness and evolutionary niche dynamics: a spatial pattern-oriented simulation experiment. *Am. Nat.*, 170, 602–616.
- Tittensor, D.P. & Worm, B. (2016). A neutral-metabolic theory of latitudinal biodiversity. *Glob. Ecol. Biogeogr.*, 25, 630–641.
- Tyler, P.A. (2003a). *Ecosystems of the deep oceans*. Elsevier, Amsterdam.
- Tyler, P.A. (2003b). The peripheral deep seas. In: *Ecosystems of the Deep Oceans* (ed. Tyler, P.A.). Elsevier Science B.V., Amsterdam, pp. 261–293.
- Webb, T.J. (2012). Marine and terrestrial ecology: unifying concepts, revealing differences. *Trends Ecol. Evol.*, 27, 535–541.
- Witman, J.D. & Roy, K. (2009). *Marine macroecology*. University of Chicago Press, Chicago.
- Yasuhara, M., Tittensor, D.P., Hillebrand, H. & Worm, B. (2017). Combining marine macroecology and palaeoecology in understanding biodiversity: microfossils as a model. *Biol. Rev.*, 92, 199–215.

CAPÍTULO 1

A GLOBAL TEST OF THE SUBSIDIZED ISLAND BIOGEOGRAPHIC HYPOTHESIS¹

¹ Menegotto A., Rangel T.F., Schrader J., Weigelt P., Kreft H. A global test of the subsidized island biogeographic hypothesis. Submitted to *Global Ecology and Biogeography*.

Abstract

The decreasing capacity of area to predict species richness on small islands (the small-island effect; SIE) seems to be one of the few exceptions to the generality of the species-area relationship. While most studies have focused on methodical aspects of how to detect the SIE, the underlying ecological factors determining this pattern remain largely unexplored. Here, we evaluate one of the few mechanisms proposed to explain the SIE, the subsidized island biogeography hypothesis, which posits that adjacent marine productivity may alter the expected species richness of small islands. We applied iterative partial regression to determine the effects of island area and marine productivity on plant species richness for 790 islands of varying sizes. We also evaluated the quadratic effect of marine productivity on residuals of the species-area relationship to test for a possible hump-shaped diversity-productivity relationship. Lastly, we employed a Geographically Weighted Regression to account for non-stationarity in the marine productivity effects. We found neither linear nor quadratic effects of marine productivity on insular species richness, at both regional and global scales. In all models, area emerged as the only predictor of plant species richness. The contribution of marine productivity was slightly apparent only in models with too few islands, but this effect was independent of island size. Our study demonstrates that marine productivity has practically no effect on plant species richness of small islands worldwide, rejecting the subsidized island biogeography hypothesis.

Keywords: Island biogeography, macroecology, small-island effect, spatial subsidy, species-area relationship, species richness

Introduction

In 1963, William Niering noted something unusual when he was studying the plant diversity of the Kapingamarangi Atoll: the number of plant species was unrelated to island area for relatively small islands (Niering 1963). This observation was intriguing because the species-area relationship is usually stated as one of the few general laws in ecology (Preston 1960; Schoener 1976; Triantis & Sfenthourakis 2012), and Niering's observation seemed a notable exception. The unusual pattern in which species richness varies independently of island area for islands below a certain size, or species richness increases at a different rate in small islands compared to larger ones, was later called the small-island effect (SIE; Heatwole & Levins 1973). However, the SIE remained largely unexplored by most ecologists for a long time (Lomolino & Weiser 2001), with only a few studies describing or discussing the phenomenon after it was first recorded (MacArthur & Wilson 1967; Rusterholz & Howe 1979; Woodroffe 1986; Morrison 1997; see Triantis *et al.* 2006 for a detailed list). A renewed interest in the SIE emerged when Lomolino (2000) suggested that the SIE may be a ubiquitous phenomenon that reflects the sigmoidal nature of the species-area relationship and provided analytical methods to model the pattern (Lomolino & Weiser 2001). Since then, many studies found evidence for a SIE in different archipelagos worldwide (Triantis *et al.* 2006; Morrison 2014; Chisholm *et al.* 2016; Gao & Perry 2016; Wang *et al.* 2016, 2018).

However, few studies have proposed ecological mechanisms to explain why species richness may vary independently of area for small islands. The first two main propositions are related to disturbance or habitat quality. Firstly, the disturbance hypothesis suggests that small islands may be more severely affected by eventual disturbances in comparison to large islands (Whittaker 1995), promoting extinction rates on small islands independently of area. Hence, an increase in island size would not necessarily result in an increase in number of species (MacArthur & Wilson 1967). Secondly, the habitat hypothesis postulates that small

islands may be ecologically restricted to one homogeneous stressful environment, tolerable to only a few species. As islands become large enough to allow the appearance of additional habitats, the typical relationship between area and species richness should emerge (Whitehead & Jones 1969; Triantis *et al.* 2003, 2006).

A third and relatively recent hypothesis is related to the productivity of the surrounding system which, for marine islands, is marine productivity. Based on ideas about spatial subsidies proposed by Gary Polis and colleagues (Polis & Hurd 1996; Polis *et al.* 1997), Anderson and Wait (2001) proposed the subsidized island biogeography hypothesis (SIB), which assume that allochthonous resources derived from adjacent waters increase insular productivity in a way that may increase or decrease the expected species richness of subsidized islands. Because smaller islands have proportionally more area exposed to the ocean, the impact of marine subsidies should be substantial on small islands, but become negligible as island size increases (Anderson & Wait 2001).

According to Anderson & Wait (2001), nutrients of marine origin may stimulate plant growth. Initially, the higher subsidized productivity would promote an increase in plant richness, therefore decreasing the slope (z) of the species-area relationship. However, after a certain level of subsidy input, diversity may decrease as productivity increases, and the z for those islands will become steeper (see Figure 1 in Anderson & Wait 2001). That is, the relationship between productivity and diversity may not be linear, but quadratic (Tilman 1982; Rosenzweig 1992; Kassen *et al.* 2000). Thus, the direction and magnitude of how much residuals deviate from the traditional species-area relationship in small islands may depend on the amount of subsidy provided by adjacent marine productivity.

Marine subsidies are indeed an important component of island food webs (Barrett *et al.* 2005; Piovia-Scott *et al.* 2011). Many studies have demonstrated that after entering the system by shore drift or seabird guano, marine-derived resources significantly increase

primary productivity (Anderson & Polis 1999; Sanchez-Pinero & Polis 2000), as well as population density of primary and secondary consumers (Polis & Hurd 1995, 1996; Sanchez-Pinero & Polis 2000; Barrett *et al.* 2005). However, few studies have tested the SIB hypothesis and evaluated the effect of marine subsidy on insular species richness in the context of the SIE. To the best of our knowledge, only one regional study on lizard richness in the Gulf of California explicitly tested the SIB hypothesis (Barrett *et al.* 2003). It is therefore largely unknown whether marine subsidies contribute to the species richness of islands and might explain the SIE. At a global scale, for example, marine productivity is highly variable (see Fig. 1). While in some waters life abounds, others are practically oceanic deserts (Lalli & Parsons 1997). Given that the flux of subsidy depends on the productivity of the adjacent waters (the donor habitat; Polis *et al.* 1997), some islands should be naturally more subsidized than others. However, no study has yet evaluated the importance of marine primary production for insular species richness considering the wide range of oceanic productivity, nor explored the relative importance of marine subsidy compared to area as island size increases.

In this study, we tested the SIB hypothesis by evaluating the effects of island area and marine productivity on angiosperm species richness of 790 islands distributed worldwide (Fig. 1), including a large number of very small islands. Assuming that high marine productivity around the islands increases marine-derived nutrients on the islands, we used marine primary production as a surrogate for potential marine subsidy. Two diversity-productivity relationships are commonly found for plants: a linear positive relationship or a quadratic relationship (i.e., “hump-shaped”; Mittelbach *et al.* 2001). We therefore applied different techniques to evaluate if plant diversity increases in islands surrounded by highly productive waters or, alternatively, if deviations from the species-area model are associated with adjacent marine productivity (Anderson & Wait 2001; see details in the data analysis).

Our results show that despite the large variation in marine primary production around the islands, marine productivity is not a predictor of insular species richness and consequently there is little support for the SIB hypothesis.

Methods

Island flora and marine productivity

To evaluate the effects of area and marine productivity on angiosperm species richness of small islands, we retrieved records of species richness, land area and geographic coordinates for 790 islands from the GIFT database (Global Inventory of Floras and Traits; Weigelt *et al.* 2019). GIFT compiles regional plant checklists and environmental data with a global coverage for macroecological studies. Our dataset encompasses islands ranging from as small as 4×10^{-6} to 102,000 km² (mean $\pm$ SD: $570 \pm 5,264$ km²), including a great number of very small islands (i.e., 420 islands had an area < 1 km²; a map with spatial distribution of island size categories can be found in the Supporting Information Appendix S1) as well as islands without any angiosperm species ($n = 50$).

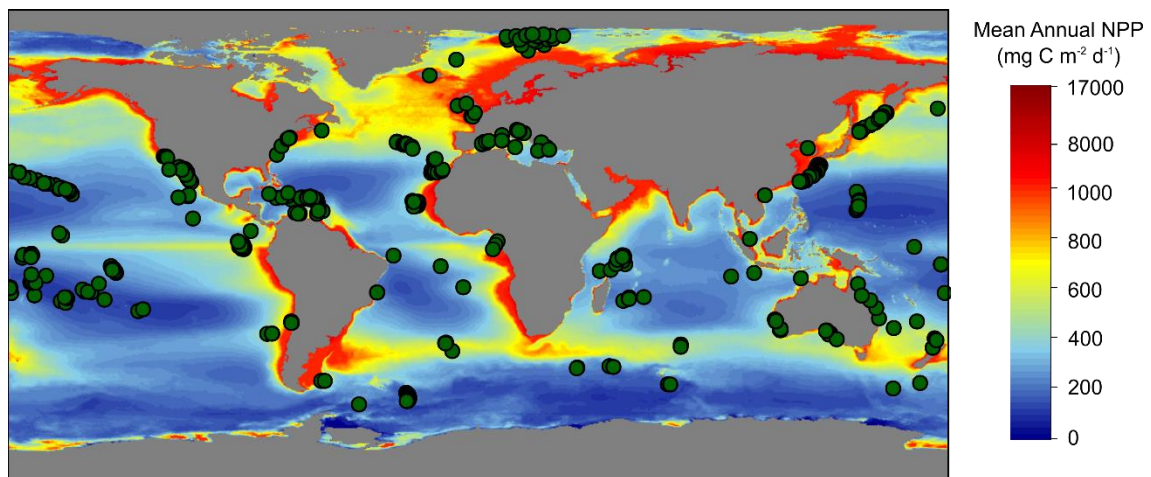


Figure 1. Global map of mean annual net primary productivity (NPP; see <http://www.science.oregonstate.edu/ocean.productivity>) in the ocean showing the spatial distribution of all the 790 islands used in this study (green circles).

For the estimation of oceanic productivity, we calculated the mean annual net primary production (NPP; $\text{mg C m}^{-2} \text{ d}^{-1}$) across the years 2003 to 2017. NPP is a function of satellite-derived chlorophyll concentration, and was generated by a vertically generalized production model (VGPM model; Behrenfeld & Falkowski 1997). Data were obtained from Oregon University database (<http://www.science.oregonstate.edu/ocean.productivity>) at the one-sixth-degree cell resolution and averaged within 0.5° cells. We considered 0.5° a reasonable resolution based on the mean foraging range of some seabirds (Thaxter *et al.* 2012; Soanes *et al.* 2016), the main subsidy vector. In addition, preliminary analyses showed that using a higher resolution (0.08°) or different NPP estimates (e.g., *Eppley*-VGPM model) did not alter the main results (Supporting Information Appendix S2). We considered the estimated NPP value of the cell where the island is located as the adjacent marine productivity around each island. When the island encompassed two or more cells, we averaged NPP values from all cells surrounding the island. If the entire island was inside a single cell without any NPP information, we attributed the NPP value of the nearest cell with available estimate of marine productivity. Overall, marine productivity around islands ranged from 117 to $4,318 \text{ mg C m}^{-2} \text{ d}^{-1}$, with a mean ($\pm \text{SD}$) of $765 \pm 549 \text{ mg C m}^{-2} \text{ d}^{-1}$.

Data analyses

In order to test the SIB hypothesis, we first used multiple regression with variation partitioning to evaluate the linear influence of island area and marine NPP on the variation in angiosperm species richness of small islands. However, determining the size under which an island can be considered small is not straightforward. Thresholds to delimit the upper limit of the SIE and, consequently, the set of islands considered small, are highly variable in the literature (Anderson & Wait 2001; Lomolino & Weiser 2001; Matthews *et al.* 2014). Likewise, the estimated threshold may vary largely for the same dataset according to the

method applied (e.g., continuous and shallow-step piecewise regression; see Supporting Information Appendix S3). Another possibility could be to explore the interaction term between area and marine NPP. However, the interaction term is highly correlated with area ($r = 0.89$; Supporting Information Appendix S4), indicating that collinearity could influence the analysis. To overcome this problem, we conducted the variance partitioning analysis in an iterative manner, instead of using an arbitrary threshold (e.g., 3 km²). We selected the 40 smallest islands in our dataset (all smaller than 10⁻⁴ km²) to start the variance partitioning. Then we applied a sequential inclusion of islands, from the smallest to the largest (following Triantis *et al.* 2006), estimating at each step the variance component explained by each predictor variable. With this method, we could observe the estimated contribution of island area and marine NPP to explaining insular species richness with increasing island size. We also used our iterative procedure to test the hypothesis that species richness in small islands is unimodally related to marine subsidies. However, here used a quadratic regression to explore the relationship between marine NPP and the residuals of the fitted species-area relationship, modelling marine NPP as a quadratic function, instead of linear. In both analyses we expected that, as the average island size increases, the effect of subsidies becomes less important, so that marine NPP becomes a poor predictor of the species richness (Fig. 2). This successive inclusion of data not only increases the mean size of the islands but also the number of islands (sample size) at each step. To evaluate the consistency of results, we therefore also conducted the same procedure in a reverse order, from the biggest to the smallest.

Another way to evaluate a potential quadratic relationship between productivity and diversity is to explore the effects of marine productivity on small island species richness at a regional scale. For example, if factors such as islands type (e.g., continental, volcanic or coral islands) or macroclimatic conditions determine island *in situ* productivity, then species

richness of different archipelagos could show opposing responses to the amount of marine resources input (Anderson & Wait 2001). In this case, the subsidy effects at a global scale should not be linear but spatially heterogeneous (i.e., non-stationary). Therefore, we assessed the spatial heterogeneity in the relative contribution of area and marine NPP to insular species richness using a variation partitioning integrated into a Geographically Weighted Regression (GWR; Fotheringham *et al.* 2002). As in a two-dimensional spatial moving window, in GWR one defines the radius of a circle (bandwidth) around a focal point, and perform a regression using the subset of the sampling data that lies within the region, assessing weights to individual samples as a function of geographical distance to the focal point. The analysis then proceeds iteratively by moving the focal point to each sample (here island) in the dataset. We defined a region with a radius 750 km around each island (our focal point) based on the mean of the maximum distance among islands within the archipelagos in our dataset (see the spatial distribution of the archipelagos in the Supporting Information Appendix S1), assigning equal weight to all islands within the radius. We performed the explained variance partitioning for each regional regression, and mapped the components across the domain (Gouveia *et al.* 2013).

Additionally, other island characteristics besides area, such as climate and isolation, are known to contribute to insular species richness (Kreft *et al.* 2008). To evaluate if the regional variation in these characteristics could increase the influence of marine NPP on species richness, we fitted a generalized linear mixed model (GLMM) with island area, distance to the nearest mainland (Weigelt & Kreft 2013), connection to the mainland during the last glacial maximum (Weigelt *et al.* 2013), temperature, precipitation, and marine NPP as the explanatory fixed variables and archipelago as the random variable, to account for the spatial organization of islands in archipelagos with distinct biogeographic histories (Bunnefeld & Phillimore 2012). Distance to the coast and connection to the mainland during

the last glacial maximum were retrieved from the GIFT database (Weigelt *et al.* 2019). Mean annual temperature and annual precipitation were obtained from CSSM simulations of modern climate models retrieved from the ecoClimate database (Lima-Ribeiro *et al.* 2015).

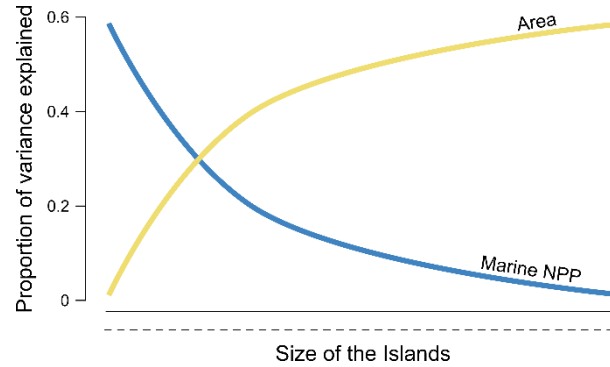


Figure 2. Hypothesized results according to the subsidized island biogeography hypothesis. Assuming a linear or quadratic relationship between productivity and diversity, variance in angiosperm species richness explained by spatial subsidies (marine NPP) should be highest for small islands, when the fraction explained by area is low, and decreases as the size of the islands and importance of area increases.

All analyses were conducted using two different transformations of species richness (S) and area (A): semi-log (S versus $\log A$) and log-log ($\log S$ versus $\log A$). Even though the SIB hypothesis has been proposed upon a log-log transformation (Anderson & Wait 2001), some studies have also found evidences of small-island effect using a semi-log transformation (Morrison 2014; Wang *et al.* 2016). Therefore, if marine subsidy has any influence on species richness of small islands, our chances to detect the pattern were not restricted to kind of data transformation. To account for the potential dependence of a biological subsidy vector, we repeated all the above analyses considering only those islands with seabirds present ($n = 377$), based on more than 60 million records of seabird occurrence from the Global Biodiversity Information Facility (<http://www.gbif.org>; see details in Supporting Information Appendix S5). We also repeated the analyses excluding all empty islands (i.e., without angiosperms; $n = 740$) and considering only those islands where both

angiosperms and seabirds were present ($n = 367$). All analyses were conducted in R version 3.5.0 (R Core Team 2018) using the packages *vegan* (Oksanen *et al.* 2018), *nlme* (Pinheiro *et al.* 2018), *raster* (Hijmans 2018) and *geosphere* (Hijmans 2017).

Results

The relative importance of island area and adjacent marine productivity for angiosperms species richness, as measured by our iterative variation partitioning, varied widely with islands size. Under both data transformations, the fraction of variance in angiosperms species richness explained exclusively by area was low and highly variable for very small islands (below 20% for islands smaller than 0.004 km^2) but increased gradually with the addition of bigger islands (maximum of 34% and 58% for semi-log and log-log transformations, respectively; Fig. 3). Conversely, the fraction explained exclusively by marine NPP was generally low but highest for small islands (maximum 6%) and approached zero after inclusion of islands larger than 0.004 km^2 . Note, however, that the fraction explained exclusively by marine NPP was always lower than the fraction explained by area (Fig. 3). In addition, a similar pattern emerged when we repeated the same procedure in reversed order, especially for the log-log transformed data; the fraction explained by area was also small and highly variable when we started the analysis with few large islands. Likewise, the fraction explained exclusively by marine NPP was highest for the biggest islands and decreased rapidly with the inclusion of other islands (Fig. 3). We found a similar result when investigating the quadratic term of marine NPP on residuals of the fitted species-area relationship. That is, only a small fraction of variation in species richness was explained by marine NPP, and this fraction was observed only when the number, not the size, of islands was small (Fig. 4).

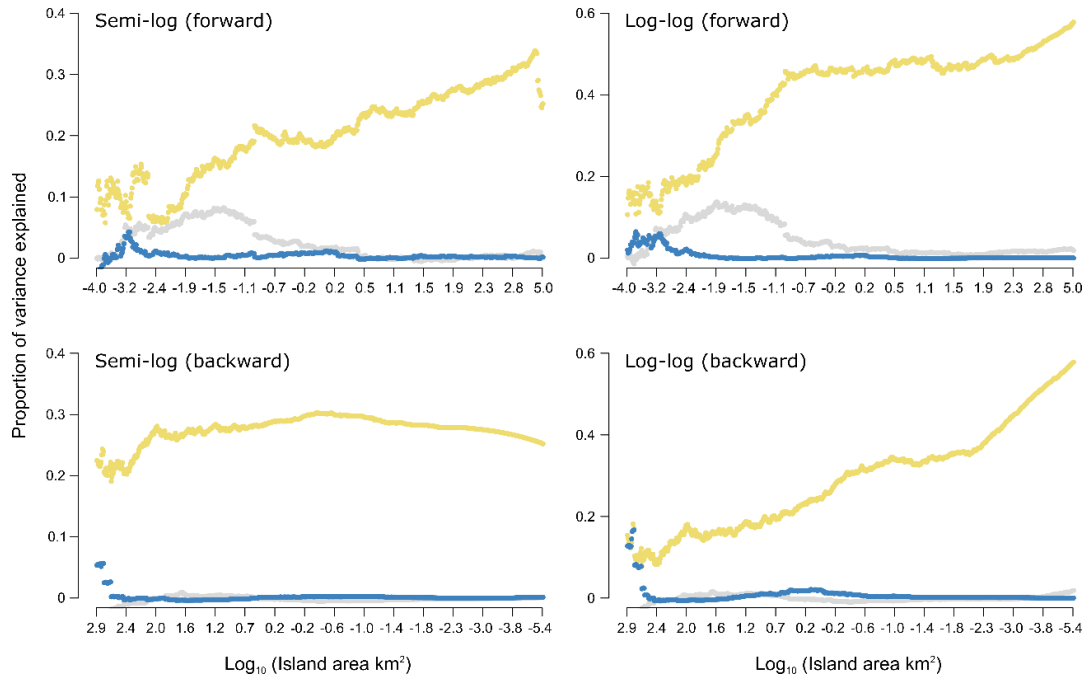


Figure 3. Variance in angiosperm species richness explained by island area (yellow dots) and mean annual marine NPP ($\text{mg C m}^{-2} \text{ d}^{-1}$; blue dots) in a multiple linear regression. Plots show the exclusive contribution of each explanatory variable and the combined fraction (grey dots) as the size of the islands increases (top) or decreases (bottom; see Fig. 2). Each tick in the x-axis represents the addition of 50 islands.

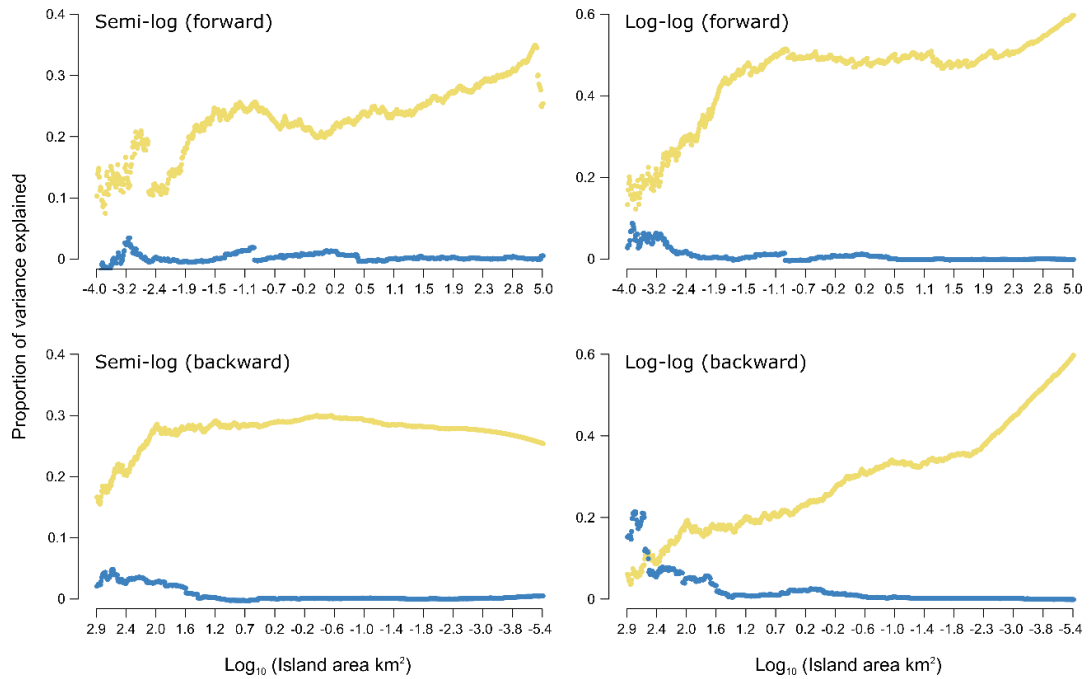


Figure 4. Variance in angiosperm species richness explained by island area (yellow dots) and mean annual marine NPP ($\text{mg C m}^{-2} \text{ d}^{-1}$; blue dots). Plots show the linear contribution of area and the quadratic effect of marine NPP on the residuals of the species-area relationship for insular angiosperms as the size of the islands increases (top) or decreases (bottom; see Fig. 2). Each tick in the x-axis represents the addition of 50 islands.

Variation partitioning based on GWR provided a spatially explicit visualization of the relative contribution of island area and marine productivity for angiosperm species richness in a regional scale. Maps show that the fraction explained exclusively by area was consistently high across the globe, varying between 40% and 90%. In contrast, the fraction explained exclusively by marine NPP was consistently low, generally below 10%. Variability in the relative contribution of both area and marine NPP seemed to increase when the number of islands in the region, and consequently in the regression, was reduced (Fig. 5; for semi-log results see Supporting Information Appendix S6). The GLMM results considering the variables associated with climate and isolation also showed no significant effect of marine NPP (Supporting Information Appendix S7).

Results were qualitatively similar when we repeated the analyses considering only the data from those islands with confirmed presence of seabirds (Supporting Information Appendix S5), island with at least one angiosperm species (Supporting Information Appendix S8) or island with at least one angiosperm species and presence of seabirds (Supporting Information Appendix S9).

Discussion

The subsidized island biogeography hypothesis, one of the few mechanisms proposed to explain why species richness on small islands may vary independently of area, predicts that resources derived from highly productive surrounding systems (i.e., the subsidies) should elevate insular productivity and alter the expected species-area relationship (Anderson & Wait 2001; Barrett *et al.* 2003). We evaluate this hypothesis by analyzing angiosperms diversity of 790 islands distributed worldwide. Our analyses involved data of 420 islands with an area $< 1 \text{ km}^2$, surrounded by waters varying mean annual productivity by two orders of magnitude from 162 to 3,547 $\text{mg C m}^{-2} \text{ d}^{-1}$. Contrary to expectations, we

did not find evidence to support the hypothesis that adjacent marine productivity may influence insular species richness.

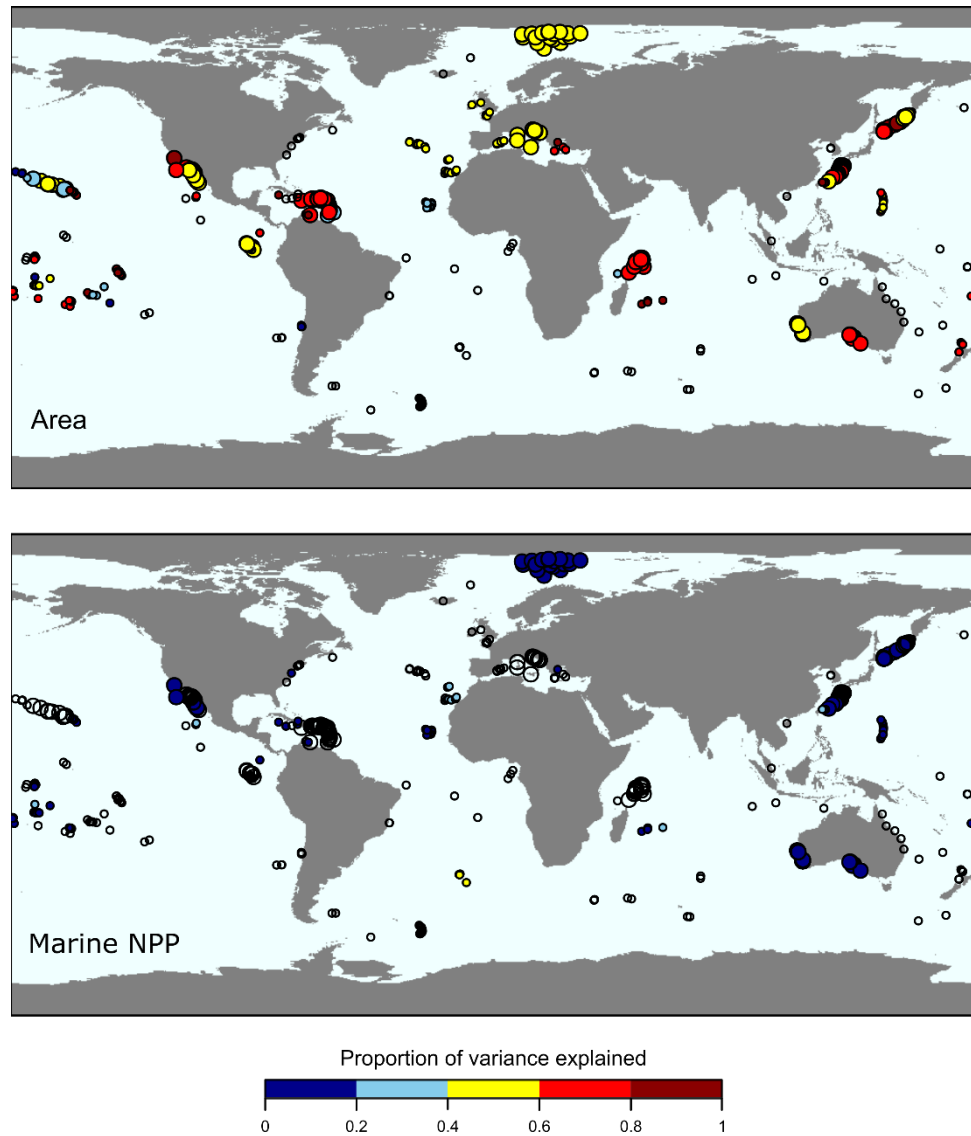


Figure 5. Non-stationarity in the effects of island area and marine NPP on angiosperms species richness. The maps show the spatial distribution of the fraction explained exclusively by area and marine NPP for each set of islands. Open circles represent non-estimated or abnormal (negative) R-squared values. To evaluate the influence of low sampling size, we plotted with small circles results of regressions calculated with less the 15 islands. Variation partitioning calculated with log-log transformed data. For results with semi-log transformation see Supporting Information Appendix S6.

Previous estimates show that marine resources may indeed dominate the productivity of islands smaller than 1 km², and its relative importance tends to decrease as island size

increases (Polis & Hurd 1996). However, the subsidy input does not seem have any effect on the number of species. Our results show that the variation in species richness explained by adjacent marine productivity was always very low (generally near zero), and never higher than area in our analyses. As in a previous study with worldwide floras, area was the strongest determinant of insular diversity (Kreft *et al.* 2008). At first, our results suggest that marine productivity has a small effect on species richness for very small islands, for which variation in species richness explained by area is low. However, our sensitivity analysis indicates that this result may be an effect of low sampling size. Alternatively, the poor predictive ability of NPP could result from the nonlinearity of the productivity-diversity relationship. For example, whereas higher productivity could lead to higher population densities and lower extinction rates, potentially increasing diversity, even higher productivity might lead to stronger competitive dominance by a few species that outcompete weak competitors, decreasing diversity (Tilman 1982; Anderson & Wait 2001; Mittelbach *et al.* 2001). However, even assuming a quadratic relationship between productivity and diversity, marine NPP was not a good predictor as would be expected if marine subsidy was affecting small islands species richness.

Despite the large variation in marine productivity around the islands analyzed here, we did not evaluate the total productivity of islands, that is, the allochthonous plus the autochthonous resources (which in this case should be the marine-derived nutrients plus the remineralization of *in situ* production). Assuming that on small islands the total productivity is dominated by marine-derived resources (Polis & Hurd 1996), variation in autochthonous productivity should have minimal effect in our analyses. Nevertheless, to ensure that the potential spatial variation in island productivity among regions (difference in temperature and precipitation, for example; Wright 1983) did not confound the interpretation of our global results, we applied methods that account for non-stationarity in the explored

relationships. For example, following the SIB hypothesis, in regions where islands have low *in situ* productivity the increase of adjacent marine production could have a positive effect on small islands species richness. Conversely, in regions where island productivity is already high, the increase of marine resources could decrease insular diversity (Anderson & Wait 2001; Mittelbach *et al.* 2001). However, even considering the possibility of a non-stationarity in the process, the fraction of variation in insular species richness explained by marine productivity was very low. The relative importance of marine productivity only increased slightly when there were too few islands in the region, suggesting, once again, a statistical artefact. Accordingly, our GLMM results considering random variation among archipelagos showed no significant effect of marine NPP on insular species richness.

Many different strategies have been used for detecting the SIE, such as different data transformations (Burns *et al.* 2009; Morrison 2014; Gao & Perry 2016; Wang *et al.* 2016), different mathematical models (Lomolino & Weiser 2001; Dengler 2010; Chisholm *et al.* 2016) and inclusion or exclusion of empty islands (Dengler 2010; Wang *et al.* 2015, 2016). Although it was not our goal to detect the SIE, we considered different alternative methods here when evaluating the importance of marine productivity for angiosperm species richness of small islands. However, there was no evidence for an influence of marine productivity on insular species richness. In addition, our analyses also included the presence of a subsidy vector, which, although not related to the SIE, is intimately associated with subsidy dynamics (Polis *et al.* 1997). In an island environment, marine resources may enter the system by shore drift or via seabird transport (Polis & Hurd 1995, 1996; Barrett *et al.* 2005). For plants, specifically, seabirds are a key subsidy vector. Previous studies have shown that the guano of seabirds, in fact, determines differences in plant growth and vegetal cover among islands (Anderson & Polis 1999; Sanchez-Pinero & Polis 2000). However, we also did not find evidence for marine productivity influencing angiosperm species richness when

we restricted our analyses to those islands where seabirds are present. Contrary to our results, studies realized in the Gulf of California islands showed that although intense guano deposition stimulates plants biomass (Polis *et al.* 1997; Anderson & Polis 1999; Sanchez-Pinero & Polis 2000; Wait *et al.* 2005), islands with seabird colonies have a species richness 2.5 times lower than those where colonies are absent (Wait *et al.* 2005), supporting the predicted descending section of the diversity-productivity hump-shaped relationship. Therefore, it is possible that the density of seabirds, the subsidy vector, is a better predictor of angiosperm diversity in small islands than just marine productivity.

Overall, our study demonstrates that the variation in species richness on small islands is not strongly related to adjacent marine productivity, at least not for plants. Future studies aiming to evaluate the effectiveness of the subsidized island biogeography hypothesis for plants should consider the influence of seabird density or colonies presence on species richness. Small-scale studies have already suggested that subsidy transport may have more importance for resource exchange than just productivity variation (Marczak *et al.* 2007; Domingos & Lana 2017). Alternatively, marine productivity may be more important for other organisms which are more dependent on shore drift (a passive subsidy vector) than on seabird presence or abundance. In fact, lizard species richness of some small islands surrounded by highly productive waters seems to be higher than expected (Barrett *et al.* 2003). However, to effectively evaluate the influence of marine productivity on insular species richness, further studies should compare these results with those from small islands surrounded by oligotrophic waters. Although many studies have already shown that marine subsidies have great influence on animal abundance (Polis & Hurd 1995, 1996; Sanchez-Pinero & Polis 2000; Barrett *et al.* 2005), the importance for animal species richness still remains to be assessed.

Acknowledgements

A.M. is supported by a Ph.D. scholarship provided by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. T.F.R. has been continuously supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, grant PQ309550/2015-7). This project is also supported by INCT in Ecology, Evolution and Biodiversity Conservation, funded by MCTIC/CNPq (grant 465610/2014-5) and FAPEG (grant 201810267000023). H.K. acknowledges funding from the DFG (grant FOR 2716 DynaCom).

References

- Anderson, W.B. & Polis, G.A. (1999). Nutrient fluxes from water to land: seabirds affect plant nutrient status on Gulf of California islands. *Oecologia*, 118, 324–332.
- Anderson, W.B. & Wait, D.A. (2001). Subsidized island biogeography hypothesis: another new twist on an old theory. *Ecol. Lett.*, 4, 289–291.
- Barrett, K., Anderson, W.B., Wait, D.A., Grismer, L.L., Polis, G.A. & Rose, M.D. (2005). Marine subsidies alter the diet and abundance of insular and coastal lizard populations. *Oikos*, 109, 145–153.
- Barrett, K., Wait, D.A. & Anderson, W.B. (2003). Small island biogeography in the Gulf of California: lizards, the subsidized island biogeography hypothesis, and the small island effect. *J. Biogeogr.*, 30, 1575–1581.
- Behrenfeld, M.J. & Falkowski, P.G. (1997). Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnol. Oceanogr.*, 42, 1–20.
- Bunnefeld, N. & Phillimore, A.B. (2012). Island, archipelago and taxon effects: mixed models as a means of dealing with the imperfect design of nature's experiments. *Ecography*, 35, 15–22.
- Burns, K.C., Mchardy, R.P. & Pledger, S. (2009). The small-island effect: fact or artefact? *Ecography*, 32, 269–276.
- Chisholm, R.A., Fung, T., Chimalakonda, D. & O'Dwyer, J.P. (2016). Maintenance of biodiversity on islands. *Proc. R. Soc. B Biol. Sci.*, 283, 20160102.

- Dengler, J. (2010). Robust methods for detecting a small island effect. *Divers. Distrib.*, 16, 256–266.
- Domingos, A.M. & Lana, P.C. (2017). Detecting multiple states of trophic connectivity between mangroves and salt marshes. *Ecosystems*, 20, 1179–1189.
- Fotheringham, A.S., Brunsdon, C. & Charlton, M. (2002). *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons, West Sussex.
- Gao, D. & Perry, G. (2016). Detecting the small island effect and nestedness of herpetofauna of the West Indies. *Ecol. Evol.*, 6, 5390–5403.
- Gouveia, S.F., Hortal, J., Cassemiro, F.A.S., Rangel, T.F. & Diniz-Filho, J.A.F. (2013). Nonstationary effects of productivity, seasonality, and historical climate changes on global amphibian diversity. *Ecography*, 36, 104–113.
- Heatwole, H. & Levins, R. (1973). Biogeography of the Puerto Rican Bank: species-turnover on a small cay, Cayo Ahogado. *Ecology*, 54, 1042–1055.
- Hijmans, R.J. (2017). Geosphere: spherical trigonometry. R package version 1.5-7. Available at: <https://CRAN.R-project.org/package=geosphere>
- Hijmans, R.J. (2018). raster: geographic data analysis and modeling. R package version 2.8-4. Available at: <https://CRAN.R-project.org/package=raster>
- Kassen, R., Buckling, A., Bell, G. & Ralney, P.B. (2000). Diversity peaks at intermediate productivity in a laboratory microcosm. *Nature*, 406, 508–512.
- Kreft, H., Jetz, W., Mutke, J., Kier, G. & Barthlott, W. (2008). Global diversity of island floras from a macroecological perspective. *Ecol. Lett.*, 11, 116–127.
- Lalli, C.M. & Parsons, T.R. (1997). *Biological oceanography: an introduction*. 2nd edn. Elsevier Butterworth-Heinemann, Oxford.
- Lima-Ribeiro, M.S., Varela, S., González-Hernández, J., Oliveira, G., Diniz-Filho, J.A.F. & Terribile, L.C. (2015). EcoClimate: a database of climate data from multiple models for past, present, and future for macroecologists and biogeographers. *Biodivers. Informatics*, 10, 1–21.
- Lomolino, M. V. (2000). Ecology's most general, yet protean pattern: the species-area relationship. *J. Biogeogr.*, 27, 17–26.
- Lomolino, M. V. & Weiser, M.D. (2001). Towards a more general species-area relationship: diversity on all islands, great and small. *J. Biogeogr.*, 28, 431–445.

- MacArthur, R.H. & Wilson, E. O. (1967). *The theory of island biogeography*. Princeton University Press, Princeton.
- Marczak, L.B., Thompson, R.M. & Richardson, J.S. (2007). Meta-analysis: trophic level, habitat, and productivity shape the food web effects of resource subsidies. *Ecology*, 88, 140–148.
- Matthews, T.J., Steinbauer, M.J., Tzirkalli, E., Triantis, K.A. & Whittaker, R.J. (2014). Thresholds and the species-area relationship: a synthetic analysis of habitat island datasets. *J. Biogeogr.*, 41, 1018–1028.
- Mittelbach, G.G., Steiner, C.F., Scheiner, S.M., Gross, K.L., Reynolds, H.L., Waide, R.B., *et al.* (2001). What is the observed relationship between species richness and productivity? *Ecology*, 82, 2381–2396.
- Morrison, L.W. (1997). The insular biogeography of small Bahamian Cays. *Ecology*, 85, 441–454.
- Morrison, L.W. (2014). The small-island effect: empty islands, temporal variability and the importance of species composition. *J. Biogeogr.*, 41, 1007–1017.
- Niering, W.A. (1963). Terrestrial ecology of Kapingamarangi Atoll, Caroline Islands. *Ecol. Monogr.*, 33, 131–160.
- Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., *et al.* (2018). vegan: community ecology package. R package version 2.5-3. Available at: <https://CRAN.R-project.org/package=vegan>
- Pinheiro, J., Bates, D., DebRoy, S. & Sarkar, D. (2018). nlme: linear and nonlinear mixed effects models. R package version 3.1-137. Available at: <https://CRAN.R-project.org/package=nlme>
- Piovia-Scott, J., Spiller, D.A. & Schoener, T.W. (2011). Predation in an island food web. *Science*, 331, 461–463.
- Polis, G.A., Anderson, W.B. & Holt, R.D. (1997). Toward an integration of landscape and food web ecology: the dynamics of spatially subsidized food webs. *Annu. Rev. Ecol. Syst.*, 28, 289–316.
- Polis, G.A. & Hurd, S.D. (1995). Extraordinarily high spider densities on islands: flow of energy from the marine to terrestrial food webs and the absence of predation. *Proc. Natl. Acad. Sci.*, 92, 4382–4386.
- Polis, G.A. & Hurd, S.D. (1996). Linking marine and terrestrial food webs: allochthonous input from the ocean supports high secondary productivity on small islands and coastal land communities. *Am. Nat.*, 147, 396–423.

- Preston, F.W. (1960). Time and space and the variation of species. *Ecology*, 41, 611–627.
- R Development Core Team (2018). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rosenzweig, M.L. (1992). Species diversity gradients: we know more and less than we thought. *J. Mammology*, 73, 715–730.
- Rusterholz, K.A. & Howe, R.W. (1979). Species-area relations of birds on small islands in a Minnesota lake. *Evolution*, 33, 468–477.
- Sanchez-Pinero, F. & Polis, G.A. (2000). Bottom-up dynamics of allochthonous input: direct and indirect effects of seabirds on islands. *Ecology*, 81, 3117–3132.
- Schoener, T. (1976). The species-area relation within archipelagos: models and evidence from island land birds. In: *The XVI International Ornithological Congress Vol. 6*. pp. 629–642.
- Soanes, L.M., Bright, J.A., Angel, L.P., Arnould, J.P.Y., Bolton, M., Berlincourt, M., *et al.* (2016). Defining marine important bird areas: testing the foraging radius approach. *Biol. Conserv.*, 196, 69–79.
- Thaxter, C.B., Lascelles, B., Sugar, K., Cook, A.S.C.P., Roos, S., Bolton, M., *et al.* (2012). Seabird foraging ranges as a preliminary tool for identifying candidate Marine Protected Areas. *Biol. Conserv.*, 156, 53–61.
- Tilman, D. (1982). *Resource competition and community structure*. Princeton University Press, Princeton.
- Triantis, K.A., Mylonas, M., Lika, K. & Vardinoyannis, K. (2003). A model for the species-area-habitat relationship. *J. Biogeogr.*, 30, 19–27.
- Triantis, K.A. & Sfenthourakis, S. (2012). Island biogeography is not a single-variable discipline: The small island effect debate. *Divers. Distrib.*, 18, 92–96.
- Triantis, K.A., Vardinoyannis, K., Tsolaki, E.P., Botsaris, I., Lika, K. & Mylonas, M. (2006). Re-approaching the small island effect. *J. Biogeogr.*, 33, 914–923.
- Wait, D.A., Aubrey, D.P. & Anderson, W.B. (2005). Seabird guano influences on desert islands: soil chemistry and herbaceous species richness and productivity. *J. Arid Environ.*, 60, 681–695.
- Wang, Y., Chen, C. & Millien, V. (2018). A global synthesis of the small-island effect in habitat islands. *Proc. R. Soc. B Biol. Sci.*, 285, 20181868.
- Wang, Y., Millien, V. & Ding, P. (2016). On empty islands and the small-island effect. *Glob.*

- Ecol. Biogeogr.*, 25, 1333–1345.
- Wang, Y., Wu, Q., Wang, X., Liu, C., Wu, L., Chen, C., *et al.* (2015). Small-island effect in snake communities on islands of an inundated lake: the need to include zeroes. *Basic Appl. Ecol.*, 16, 19–27.
- Weigelt, P., Jetz, W. & Kreft, H. (2013). Bioclimatic and physical characterization of the world's islands. *Proc. Natl. Acad. Sci.*, 110, 15307–15312.
- Weigelt, P., König, C. & Kreft, H. (2019). GIFT – A Global Inventory of Floras and Traits for macroecology and biogeography. *bioRxiv*, doi:10.1101/535005.
- Weigelt, P. & Kreft, H. (2013). Quantifying island isolation - insights from global patterns of insular plant species richness. *Ecography*, 36, 417–429.
- Whitehead, D.R. & Jones, C.E. (1969). Small islands and the equilibrium theory of insular biogeography. *Evolution*, 23, 171–179.
- Whittaker, R.J. (1995). Disturbed island ecology. *Trends Ecol. Evol.*, 10, 421–425.
- Woodroffe, C.D. (1986). Vascular plant species-area relationships on Nui Atoll, Tuvalu, Central Pacific: a reassessment of the small island effect. *Aust. J. Ecol.*, 11, 21–31.
- Wright, D.H. (1983). Species-energy theory: an extension of species-area theory. *Oikos*, 41, 496–506.

CAPÍTULO 2

TAKING A MARINE PERSPECTIVE ON THE LATITUDINAL DIVERSITY GRADIENT²

² Manuscrito desenvolvido durante exame de qualificação de doutorado.

Abstract

The increasing species richness from poles to equator intrigues ecologists for decades, and a large number of hypotheses has been proposed to explain this geographical pattern. However, most of these hypotheses have been developed and tested based in terrestrial data, suggesting that non-analog environmental features of the marine realm reserve great opportunity to evaluate these ideas. In this essay I realized a conceptual review of the main evolutionary hypotheses proposed to explain the origin of the latitudinal biodiversity gradient under a marine perspective. I found that hypotheses based on assumptions of physiological tolerance and dispersal ability predict a tropical and polar higher speciation rate, because both biotas are composed by stenothermal species with short (or absent) larval phase. Evolutionary speed was the only hypothesis supporting higher tropical speciation rate. Studies show that temperature indeed affects generation time and genetic evolution of marine organisms, though direct effects of temperature on speciation rate still need more evidences. The area hypothesis has the least support in empirical data, while no data was available to effectively evaluate the predicted evolutionary effects of biotic interaction. Given the recent contradiction in the predominance of tropical or polar speciation rate, it is still hard to determine which hypothesis has more empirical support. However, this exercise illustrates how a further comprehension of the geographical variation in marine speciation rate and detailed descriptions of biodiversity patterns may be essential to advance our comprehension about the origin of the latitudinal diversity gradient.

Keywords: Diversification, evolution, latitude, macroecology, marine biodiversity, speciation rate, species richness.

A challenging and fascinating mystery

The geographical pattern of increasing species richness from poles to equator, the so-called latitudinal diversity gradient, is one of the most intriguing questions in ecology. Along ago this pattern was first noted by early naturalists like Alexander von Humboldt and Alfred Wallace, and since then ecologists struggle to resolve this mystery. Many hypotheses have been proposed, tested and reviewed to explain the latitudinal variation in species richness (Pianka 1966; Rohde 1992; Willig *et al.* 2003; Mittelbach *et al.* 2007), but no consensus has emerged on this issue yet. After two centuries of discussion, the latitudinal diversity gradient remains a pattern in search of a theory (Rosenzweig 1992). But why is so difficult to determine the causal mechanism for this general pattern?

Most of the hypotheses proposed to explain the latitudinal variation in species richness suggest some environmental or geographical variable as the main driver of the observed pattern. One limitation to our advance on the issue, however, is that most of these variables are generally spatially correlated across the latitudes, at least on land. Temperature, productivity, area, climate stability, all increase toward the tropics (Rosenzweig 1992; Dynesius & Jansson 2000; Gillman *et al.* 2015), what unable us to effectively identify the primary mechanism driving species richness pattern across the latitudes. I believe that our understanding of the underlying process should certainly be improved if we look at the predicted mechanisms in a large-scale environment where the spatial configuration of hypothetical drivers (biotic or abiotic) follows a different pattern from that observed on land, as the marine realm.

In the ocean, latitudinal distribution of some environmental and biological variables is very different from what would be expected on land. The distribution of primary production, for instance, is lower at the tropics, peaking at mid-latitudes (Box 1). Although the sea surface temperature is similar to that of land, it tends to be homogeneously cold

without seasonal variation in the deep sea (Box 1). In addition, absence of major physical barriers as well as the presence of a clear spatial pattern in life history of marine organisms (Marshall *et al.* 2012) allows us to test specific assumptions about dispersal and isolation processes. Because of such particularities, terrestrial and marine realms may be seen as completely different systems. Yet, both systems exhibit a similar latitudinal trend in species richness variation (Hillebrand 2004a; Kinlock *et al.* 2018). Therefore, these differences in environmental variability and biological features at large scale present a major opportunity for theoretical advances in our understanding of macroecological patterns (Parmesan *et al.* 2005; Webb 2012). After all, if ecological and evolutionary mechanisms generating the latitudinal diversity gradient can be generalized, they should cross the land-sea boundary.

Despite the benefits of cross-domain collaboration in theories and methods, testing hypotheses and statistical tools developed in one system in another, terrestrial and marine ecology have developed a large gap in communication (Steele 1991; Menge *et al.* 2009; Webb 2012). Terrestrial ecologists are little educated to think about marine environments (e.g., few marine examples in ecological textbooks, few presentations placed in specific sections in ecological meetings; Munguia & Ojanguren 2015) and usually ignore most of the aquatic literature, even when published in “general” ecological journals (Menge *et al.* 2009). On the other hand, marine ecologists usually do not place their work in a broader context (Munguia & Ojanguren 2015), limiting their potential for testing hypotheses based on terrestrial data. To illustrate how cross-domain collaboration could advance our understanding about the causes of the latitudinal gradient of biodiversity, in this essay I will briefly review the history of marine research on the subject and then evaluate conceptually the main hypotheses proposed to explain this old mystery under a marine perspective.

A brief history of the latitudinal diversity gradient in the sea

The first description of a latitudinal diversity gradient for marine species is attributed to the Danish zoologist Gunnar Thorson, who demonstrated the increasing number of hard-bottom species (crustacean and mollusk) toward the equator (Thorson 1957). A few years later, the same pattern was described for other benthic species of shallow waters as gastropods (Fischer 1960) and bivalves (Stehli *et al.* 1967), as well as for pelagic species from plankton (foraminifera; Stehli *et al.* 1972) and nekton (fishes; Rohde 1978). Surprisingly, further samplings on deep waters revealed the existence of a latitudinal pattern even in the relatively homogeneous deep sea (Rex *et al.* 1993). The number of studies was still small in this period and highly restrict to the North Atlantic, so that the generalization of a global pattern comparable to that seen on land was criticized (Clarke 1992). Some studies even questioned the existence of a latitudinal gradient of species richness in the less sampled southern hemisphere (Rex *et al.* 1993; Gray 2002). However, the number of studies increased after 2000s, and today the latitudinal gradient of marine biodiversity is a fact (Hillebrand 2004a, b; Kinlock *et al.* 2018). Despite some exceptions (e.g., algae, mammals and birds), the latitudinal gradient of marine biodiversity had been described as a unimodal pattern until recently. Studies published in the last years, however, have proposed a bimodal pattern peaking at mid-latitudes (Chaudhary *et al.* 2016), at least in the oceanic zone (Tittensor *et al.* 2010) or the deep sea (Woolley *et al.* 2016). Such contrasting patterns reveal the necessity for a better understanding of the marine latitudinal gradient but also reveal its potential for exploring causal mechanisms.

For the first marine macroecologists, probably influenced by Humboldt, Wallace and Dobzhansky, latitudinal changes in species richness clearly resulted from evolutionary processes associated with gradients in environmental stability. Both Fischer (1960) and Sanders (1968) believed that climatic fluctuations at high latitudes should limit species

diversity due to either periodical extinctions or reduced time for ecological specializations, decreasing speciation rate. Conversely, the long-time environmental stability of the warm tropics should favor the fast and continuous evolution of mature communities, increasing speciation events. Posterior studies with marine fossils showed that speciation events were in fact higher in the tropics (Stehli *et al.* 1969), although there was no clue about the possible causes of such variation in evolutionary rate. In the 90's evolutionary and ecological questions apparently split up. Evolutionary studies started to investigate latitudinal variation in speciation and extinction rates mainly through fossils records (e.g., Jablonski 1993), though without major advances in determine the causes of such variation. Meanwhile, ecological studies focused on correlative approach that explored the relationship between richness and hypothesized causal agents (e.g., Fraser & Currie 1996; Roy *et al.* 1998), but without further discussions about evolutionary processes. Today, evolutionary and macroecological studies are joining together again, providing clues about the origin of the marine biodiversity gradient (Jablonski *et al.* 2017; Yasuhara *et al.* 2017). Summing this to the increasing number of physiological and genetic studies with marine species, we have now some empirical base to explore conceptually the gradient of biodiversity in the sea.

Hypothetical explanations under a marine perspective

The living species which exist in the world today are a product of evolution. Therefore, any explanation for the latitudinal diversity gradient requires the explicit consideration of the three macroevolutionary process that directly change the number of species: speciation, extinction, and dispersal (Ricklefs 2004; Goldberg *et al.* 2005). Some hypotheses suggest that these processes do not necessarily change geographically. Instead, tropical latitudes accumulate more species only because they are old and few species can make the transition to temperate climates (e.g., the tropical niche conservatism hypothesis;

Wiens & Donoghue 2004). However, many alternative models suggest that tropical diversity may arise from geographical variation in diversification and/or dispersal rates (Mittelbach *et al.* 2007; Roy & Goldberg 2007; Fine 2015; Alves *et al.* 2017). The out of the tropics hypothesis, for example, postulate that the tropics have both a higher origination and lower extinction rate, with a positive net movement of species to extratropical latitudes (Jablonski *et al.* 2006).

The set of hypotheses which predict higher speciation rate in tropical communities to explain the gradient of biodiversity have had historically broad support in marine data. Fossils of planktonic (Allen & Gillooly 2006) and benthic species (Hecht & Agan 1972; Jablonski 1993; Jablonski *et al.* 2006; Martin *et al.* 2007; Powell 2007) as well as phylogenetic analysis of reef fishes (Siqueira *et al.* 2016) all suggest that speciation rate is higher in tropical waters. However, recent phylogenetic studies have shown that speciation rate may actually be higher in polar latitudes, even when the diversity of the taxa peaks in the tropics (Rabosky *et al.* 2018; O'Hara *et al.* 2019). The reasons why speciation rate might differ between tropical and extratropical latitudes are still not completely understood. Some mechanisms have been proposed to explain the higher speciation rate in the tropics (reviewed by Mittelbach *et al.* 2007; Fine 2015). But, to the extension of my knowledge, they have never been discussed deeply under a marine perspective, even less considering how these same mechanisms could lead to a higher speciation rate in the cold polar seas. Given the evidences cited above, to explore the hypothetical causes behind the evolutionary cradles in the sea may serve a good exercise to understand the global pattern of biodiversity.

In this section I will review the main mechanisms proposed to explain the latitudinal variation in diversification rate, focusing in evaluate their assumptions and predictions for a higher origination rate against environmental and biological features of the marine realm. I am assuming here that marine species diversity (not necessarily speciation) increases toward

the tropics, which has been supported by different meta-analyses (Hillebrand 2004a; Kinlock *et al.* 2018). Moreover, that seems to be the common pattern observed in the continental shelf (Tittensor *et al.* 2010; Woolley *et al.* 2016; O'Hara *et al.* 2019), from where comes most part of the ecological, physiological, genetic and paleontological data in marine biodiversity. Nevertheless, I acknowledge that the bimodal pattern may be relevant as well and I will discuss it when appropriated.

Geographical area

The hypothetical mechanism by which area would affect speciation rate predicts that larger areas should support larger range sizes, turning species more exposed to ecological heterogeneity, physical barriers or peripheral isolation (Terborgh 1973; Rosenzweig 1992), feeding allopatric speciation. Oceanic area in fact is larger at the tropics (Allen & Gillooly 2006). Accordingly, a positive area effect on speciation rate and species richness has been supported by neutral simulations (Tittensor & Worm 2016) and observed in planktonic species (Allen & Gillooly 2006). However, the speciation rate of planktonic species exhibits a tropical peak even after controlling for area (Allen & Gillooly 2006). Moreover, fossil analyses revealed that, at least for gastropods, large range does not lead to either higher speciation rate or total production of species (Jablonski & Roy 2003). Conversely, large range size seems reduce speciation rate (Jablonski 1986).

Even if not a source of origination, larger areas could accumulate more species simply because less extinction events would occur in larger communities (Rosenzweig 1992). In this case, larger areas should harbor higher diversity (e.g., reef fish; Barneche *et al.* 2019). However, it may not be so common as would expected (Rohde 1992). Along the coast of the Americas, for example, continental shelf area increases with latitude (Roy *et al.* 1998) whereas benthic species richness follows an inverse pattern (Steele 1988; Roy *et al.*

1998, 2000; Macpherson 2002; Astorga *et al.* 2003). Although exceptions exist (Valdovinos *et al.* 2003), they seem more related to development mode than area (Roy *et al.* 2000; Fernández *et al.* 2009; see discussion in the *Climate change* hypothesis). Therefore, area size does not seem a mechanism robust enough to support neither higher speciation rate nor to explain the present gradient in species diversity.

Physiological tolerance

The tolerance of any organism must span the range of climatic conditions to which it is exposed through its life (Stevens 1989). Therefore, species exposed to high seasonal variation in climatic conditions (e.g., temperature) must have wide climatic tolerance, whereas those experiencing smaller climatic fluctuation during the year may develop greater physiological specialization. Because temperature seasonality increases with latitude, at least on land, tropical species should be more specialized and, consequently, spatially limited by their reduced physiological tolerance. This means that climatic barriers are “higher” for tropical species than for temperate ones (Janzen 1967), which should reduce their opportunity for dispersal and gene flow, enhancing genetic divergence between population and, consequently, allopatric speciation events. Marine species from highly seasonal latitudes are indeed more thermotolerant than those of relatively stable waters (Sunday *et al.* 2011), even in embryonic stages (Woolsey *et al.* 2015). However, seasonal variation in temperature does not increase linearly with latitude along the oceans as observed on land. Seasonal variation in oceanic temperature peaks at middle, temperate latitudes and declines toward the poles (Parmesan *et al.* 2005; Box 1). Accordingly, tropical and polar marine biota are mainly composed by stenothermal species, tolerating only narrow ranges of temperature (warm and cold, respectively; Pörtner *et al.* 2007; Somero 2010; Sunday *et al.* 2011; Peck *et al.* 2014). Polar species exhibit longer time of temperature acclimation than temperate ones

(Peck *et al.* 2014), with some species even unable to acclimate a few degrees (Peck *et al.* 2009). If seasonal stability and physiological specialization were the trigger for speciation, origination rate should peak in tropical as well as polar latitudes, as may be the case (O'Hara *et al.* 2019). This means that higher tropical diversity could result from dispersal and/or extinction events in high latitudes. The Achilles' heel of this hypothesis is the deep sea. As environmental conditions are relatively stable in deep waters (Box 1), this hypothesis predicts that deep-sea lineages should speciate uniformly, regardless of latitude. However, speciation rate is still faster for high-latitudes deep-sea species (Rabosky *et al.* 2018; O'Hara *et al.* 2019).

Climate change

During the history of the Earth, climate has fluctuated widely. The magnitude of climatic oscillations, however, is not spatially homogeneous. It is greater at high latitudes and decrease toward the tropics (Dynesius & Jansson 2000). The climate change hypothesis predicts that the greater climatic oscillation at high latitudes should lead to strong selection for wider environmental tolerance, generalism and dispersal ability, producing latitudinal gradients in these properties. Because dispersal ability may reduce isolation among populations, speciation rates should therefore decrease toward high latitudes (Dynesius & Jansson 2000). One could argue that the marine biodiversity gradient was established long before the greater climatic oscillation of the Quaternary (Mittelbach *et al.* 2007), as evidenced in fossil records of brachiopods, foraminifera, bivalves and other invertebrates from Paleozoic and Mesozoic (Stehli *et al.* 1969; Crame 2001; Powell 2007; Alroy *et al.* 2008). However, I will focus here on the assumptions associated with the species' biological traits.

Polar marine species are not more generalist in climatic tolerance as predicted (as showed before), while generalism in the use of resources has contradictory patterns (e.g., increase for gastropods but decrease for fishes; (González-Bergonzoni *et al.* 2012; Crame *et al.* 2018). However, the main assumption of the climate change hypothesis is that high-latitude biota should be composed by highly dispersive species, what would effectively increase gene flow and reduce speciation rate. For most marine species, especially the benthic ones, dispersal occur in larval stage, with planktonic larvae having longer dispersal distances than non-planktonic larvae (Shanks 2009). At regional scale, some studies have already shown that species with short larval duration or non-planktonic larvae display relatively strong genetic population structure in comparison to species with planktonic larval stage (Bradbury *et al.* 2008; Chust *et al.* 2016; Ellingson & Krug 2016; Modica *et al.* 2017). Consequently, those species with non-planktotrophic larvae (i.e., direct development or lecithotrophic larvae) also exhibit higher speciation rates than other species (Jablonski 1986). According to the hypothesis, we could then expect a predominance of species with dispersive planktonic larvae in high latitudes, but that is not the case. Contrary to the expected, the increasing latitude is marked by a reduction in species with planktonic larvae and a higher proportion of non-planktonic development (Fernández *et al.* 2009; Marshall *et al.* 2012; but see Leis *et al.* 2013).

Despite the latitudinal change in development mode cited above, the gradient in dispersal capacity is not necessarily linear. Although planktonic larvae predominate in lower latitudes, it is important to emphasize that larval duration is not homogenous in these species. Rather, it increases in cooler temperatures due metabolic processes, peaking at mid-latitudes (O'Connor *et al.* 2007). Indeed, higher dispersal distance and reduced genetic structure have been observed in temperate species (Bradbury *et al.* 2008; Jones *et al.* 2009). This development mode may become less common in the poles because longer larval duration in

extremely cold waters should increase mortality risk. Therefore, natural selection must favor those organisms with short or non-planktonic development modes in colder habitats (O'Connor *et al.* 2007). In a nutshell, dispersal capacity seems to be relatively reduced in both tropical and polar latitudes, supporting the hypothesis that dispersal limitation may represent an important mechanism for geographical variation in speciation rates. However, it does not seem to be triggered by climate change.

Evolutionary speed

Unlike the previous set of hypotheses, that demanded geographical isolation to increase speciation, the evolutionary speed hypothesis focuses only on the temperature effects on mutation rate. Kinetic effects of environmental temperature are proposed to increase mutation rate and genetic evolution direct or indirectly via metabolic effects, accelerating natural selection (Rohde 1992). Direct effects include by-products released by metabolism (e.g., free radicals) which can oxidize DNA and induce mutations (see details in Evans & Gaston 2005). Indirectly, temperature might increase genetic evolution by reducing generation time, increasing mutations per unit time and genetic divergence (Rohde 1992). In the marine realm, temperature seems indeed to determine generation time. Lifespan, growth rate and maturation time of marine organisms all demonstrate clear trends with latitude, between and within species (Trip *et al.* 2014; Moss *et al.* 2016), even after controlling for phylogenetic dependence (Ridgway *et al.* 2011). This pattern is consistent even in the absence of body size variation (Munch & Salinas 2009; Moss *et al.* 2016). Interestingly, the same pattern is observed along the bathymetric gradient, with long-lived, slow-growing and late maturation species occurring in deep waters (Garcia *et al.* 2008; Drazen & Haedrich 2012). Few studies have tested the temperature effect on mutation rate, but the evidences to date are consistent with the prediction that genetic evolution and speciation are generally

faster in warm waters, varying spatially both in latitude and depth (Allen *et al.* 2006; Wright *et al.* 2011). It is still not clear, however, if temperature would be affecting this genetic process direct or indirectly.

A big challenge for the evolutionary speed hypothesis is to explain the gradient of biodiversity observed in the cold and relatively homogeneous deep waters. If a small increase in temperature is enough to affect generation time or mutation rate, so the slightly higher temperature of the tropical deep sea (Box 1) could still increase tropical speciation and generate a unimodal richness pattern. Some studies have already explored the effects of small temperature variation on life history traits related to generation time. Evidences suggest that growth rate increases almost four times when temperature rise from 0° to 3°C in an Antarctic scallop (Heilmayer *et al.* 2005). Similarly, a few studies with upper slope fishes (nearly 250 – 1500 m) revealed that even at this depth populations of high latitudes seem to have late maturation, slow growth rate and longer lifespan than those of lower latitudes, where temperature is only 4° to 5°C higher (Winton *et al.* 2014; Rochowski *et al.* 2015). Obviously, given the small variation in deep waters temperature, we should expect a deep-sea biodiversity gradient less steep than that from shallow waters, as in fact was observed (Hillebrand 2004b). The main problem here is that deep-sea tropical diversity seems result from high-latitudes dispersal rather than higher rates of *in situ* speciation (O'Hara *et al.* 2019), which may not be explained by variation in temperature.

A recent review (Gillman & Wright 2014) has questioned the isolated effect of temperature on the evolutionary speed since habitats with warm dry climates are species poor. The authors proposed the “integrated evolutionary speed hypothesis”, which suggested that rates of genetic evolution are associated with temperature and water availability. Both variables could affect genetic evolution independently or via a combined effect through productivity. Experiments with an Antarctic echinoderm revealed that food supply had great

impact on metabolic rate once polar waters have no major variation in temperature (Brockington & Clarke 2001). Could, then, metabolic rate mediated by intense food availability speed genetic evolution and speciation rate in the cold polar latitudes? That is still not so clear. Mainly because another hypothesis suggests that marine speciation rate is reduced in habitats where trophic supply is highly seasonal as the poles, even if total productivity is relatively higher than other regions (Valentine & Jablonski 2015). Therefore, while gradients in energy or resource stability may explain a higher speciation rate in the tropics, it does not seem a good mechanism to explain the alternative polar cradle recently described for some groups.

Biotic interaction

One of the first hypotheses proposed by early naturalists and ecologists to explain the latitudinal gradient of biodiversity is that tropical richness results from higher biotic interaction in low latitudes (Mittelbach *et al.* 2007). The hypothesis simply assumes that tropical climate stability favors adaptations to biotic interaction, which should increase the rate of speciation. In high latitudes, on the other hand, low and variable temperature must inhibit such interactions, consequently reducing speciation events (Dobzhansky 1950; Fischer 1960; Sanders 1968). In fact, biotic interactions increase toward the tropics in marine communities (Schemske *et al.* 2009; Freestone *et al.* 2011; González-Bergonzoni *et al.* 2012; Longo *et al.* 2019; for an exception see Torchin *et al.* 2015). But latitudinal variation in biotic interaction does not represent evidence of a causal effect. Conversely, it may be just a consequence of the species richness gradient. To truly evaluate this hypothesis, we need understand its underlying evolutionary processes.

A recently proposed mechanism (Schemske 2002, 2009) assumes that under strong biotic selection, separated tropical populations are forced into new coevolutionary pathways

that increase phenotypic and genetic divergence. In contrast, temperate populations should experience the same environmental selection of severe cold climate, declining phenotypic and genetic divergence among populations. Unfortunately, there is no synthesis about latitudinal variation in nature of selection either in marine or terrestrial realm to test this hypothesis. A study with a high latitude fish (50° - 70°) revealed that separate populations in each side of the North Atlantic have the same loci under directional selection associated with temperature variation (Bradbury *et al.* 2010), reinforcing the importance of environmental factors on natural selection in high latitudes. However, once a species had passed through this large-scale environmental filter, why cannot it develop adaptations to biotic interaction? At the Antarctica, notothenioids developed antifreeze glycoproteins that enable these ancestrally benthic fishes to survive and adapt in extremely cold waters (Matschiner *et al.* 2011). Ecological opportunity allowed this environmentally adapted lineage diversify through disruptive selection along the last millions of years, adapting its cranial format to different diet items and even reinvading the pelagic zone (Hu *et al.* 2016). Similar example is found in post-glacial lakes (in the absence of more marine cases), where populations of the same fish species present distinct morphological adaptations to their specific diet and habitat, even in sympatry (Knudsen *et al.* 2011). These examples illustrate how biotic selection may also be important for high latitude species, but still do not provide any evidence about the prevalence of biotic or abiotic selection across the latitudes.

Concluding remarks

The result of this conceptual review shows that, contrary to the terrestrial system, each mechanism proposed to explain the global diversity pattern generates a different prediction for cradles distribution in the marine realm. Area was the least robust hypothesis, with little support for its prediction in both fossil and current data. Physiological tolerance

and dispersal capacity hypotheses predict similarly higher speciation rate in tropical and polar latitudes. This means that the current diversity pattern should result mainly from dispersal and extinction processes. Variation in trophic supplies, for example, would turn species subject to stochastic extinction in the poles, whereas steady primary production would favor the accumulation of high diversity in lower latitudes (Valentine & Jablonski 2015). What is still not so clear is why some taxonomic groups show higher speciation rate in the tropics while others in the poles, instead of both (but see O'Hara *et al.* 2019). Future studies should investigate if this contrasting pattern result from any intrinsic feature of each group or some methodological issue. Most results with higher tropical speciation come from fossil analyses, while higher polar speciation come from molecular studies. Further investigations encompassing different taxa and comparing results from fossil and molecular data may shed some light in the predominant pattern of speciation rate in the oceans and the accuracy of these two mechanisms.

Considering the hypothesis that the tropics are the main hotspot of species formation, evolutionary speed should be the most supported explanation in the marine realm. Higher speciation rate per unit time would not only allow species diversify in the tropics but also develop adaptations to explore new environments. According to fossil analysis, climate-niche conservatism is indeed violated by tropical clades with higher diversification (Jablonski *et al.* 2013). Even with lower thermal tolerance, higher mutation rate could give to some tropical species opportunity for adaptation to lower temperatures, which would allow them expand their range and colonize extra tropical regions, increasing the latitudinal range of some lineages as observed in the fossil records (Stehli *et al.* 1969; Jablonski *et al.* 2006; Martin *et al.* 2007).

The role played by biotic interaction in the spatial variation of speciation rate still remains unclear. Mainly because it is difficult to determine the relative importance of biotic

and environmental process driving speciation, and identify the types of selection for different species or clade is surely not a trivial task. An alternative hypothesis is that biotic interactions contribute as a positive feedback system that enhances biodiversity rather than be a primary generator of diversification gradients (Gillman & Wright 2014). For example, biotic selection could interact with higher temperature-mediated genetic evolution to produce a stronger latitudinal gradient of biodiversity (see a similar result with area in Tittensor & Worm 2016). Recent studies have already shown that biotic interaction seems interact with metabolic rate and drive the geographical diversification of marine predators (Grady *et al.* 2019).

It is noteworthy that many of the mechanisms proposed to explain higher rates of tropical speciation are grounded in the importance of allopatric events. This is interesting because the presence of few physical barriers and the potentially extensive dispersal of marine species should reduce opportunities for allopatric speciation in the ocean. Indeed, sympatric speciation seems to be the predominant speciation process in the marine realm (Norris & Hull 2012; Bowen *et al.* 2013). Another interesting point is that most (if not all) of the evolutionary hypotheses and their predictions do not support a scenario of bimodal species richness as described in recent studies (Tittensor *et al.* 2010; Chaudhary *et al.* 2016; Woolley *et al.* 2016). Therefore, a general bimodal pattern should only arise from dispersal instead of diversification process.

In summary, I show here that the study of marine biodiversity patterns may represent a key avenue to understand the global pattern in species distribution. Given the recent discoveries of high polar speciation rate, it is still hard to determine which hypothesis have more empirical support. However, the strength of each hypothesis will certainly change as more taxonomic groups have their phylogeny and fossil records investigated, as well as more physiological and functional data are compiled to confront the assumptions of each proposed

mechanism. Unfortunately, our knowledge of marine biodiversity tends to be highly limited, biased to a few well studied taxonomic groups (Reusch 2014) and geographical regions (Mora *et al.* 2008; Menegotto & Rangel 2018). Therefore, additional work and syntheses of marine biodiversity patterns may be essential to advance our comprehension about the origin of the latitudinal diversity gradient.

Box 1. A quick guide of the marine environment for macroecologists

Ocean comprehends a tridimensional environment that covers 71% of the planet. At the surface, sunlight warms up superficial waters generating a latitudinal gradient of temperature comparable to that of land (Figure 1a). Such temperature may be relatively uniform in the upper well-mixed layer (100 to 200 m), but decreases rapidly with increasing depth (the thermocline), and becomes relatively stable after 1,000 m. Below this warm upper layer, there is little latitudinal variation in temperature (Figure 1a, e). As you can imagine, photosynthesis is largely restricted to this thin upper layer where light is available, also called the euphotic zone. Nutrients in the euphotic zone are rapidly consumed and quickly become scarce. Below this layer nutrients abound, but there is no light. Nutrients and light limitation make the marine latitudinal pattern of productivity completely different from that on land (Figure 1c, d). Tropical primary production is generally reduced because the low thermal seasonality at the tropics (Figure 1b) keeps the upper warm and lower cold waters stratified. At mid-latitudes, winter cooling and storms destabilize the water stratification, what enable that nutrients return to surface and increase productivity after winter, when sunlight increases. At the poles, cold waters are not stratified and nutrients are not so limiting. However, primary production is restricted to the few weeks there is light. On average, marine productivity is higher at mid-latitudes (Figure 1c). In the deep sea, practically devoid of autotrophs (except chemosynthesis), the resources maintaining secondary production come almost entirely from organic matter sinking from productivity superficial waters.

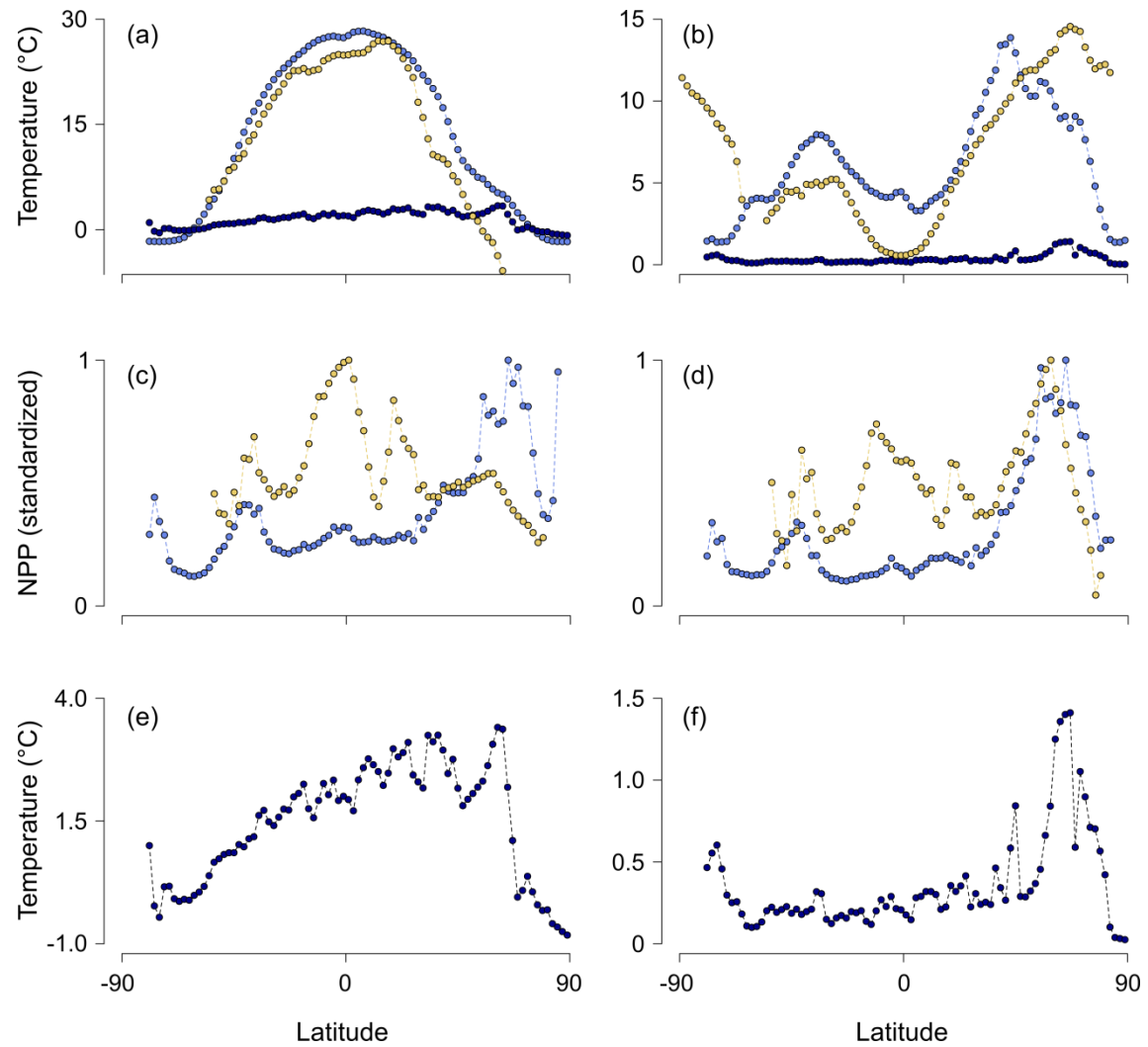


Figure I. Latitudinal distribution of temperature and productivity across the latitudes in terrestrial and marine realm. Annual mean (a) and standard deviation (b) temperature in land (yellow), sea surface (light blue) and sea bottom (below 200 m; dark blue). Note that polar land temperature goes down to -44° and -20°C in south and north poles, respectively, but these data are omitted in the figure. Annual mean (c) and standard deviation (d) net primary productivity (NPP) in land and sea surface. In panels (e) and (f) I show the same bottom temperature from panels (a) and (b) in more details. Data were retrieved from: WorldClim, Bio-ORACLE, Earth Observatory and Oregon State University. All data were converted to 2° resolution before mean calculation.

Acknowledgements

I am very grateful to Fabricio Villalobos and José Alexandre F. Diniz-Filho for valuable comments on an early version of the manuscript.

References

- Allen, A.P. & Gillooly, J.F. (2006). Assessing latitudinal gradients in speciation rates and biodiversity at the global scale. *Ecol. Lett.*, 9, 947–954.
- Allen, A.P., Gillooly, J.F., Savage, V.M. & Brown, J.H. (2006). Kinetic effects of temperature on rates of genetic divergence and speciation. *Proc. Natl. Acad. Sci.*, 103, 9130–9135.
- Alroy, J., Aberhan, M., Bottjer, D.J., Foote, M., Fursich, F.T., Harries, P.J., *et al.* (2008). Phanerozoic trends in the global diversity of marine invertebrates. *Science*, 321, 97–100.
- Alves, D.M.C.C., Diniz-Filho, J.A.F. & Villalobos, F. (2017). Geographical diversification and the effect of model and data inadequacies: the bat diversity gradient as a case study. *Biol. J. Linn. Soc.*, 121, 894–906.
- Astorga, A., Fernández, M., Boschi, E.E. & Lagos, N. (2003). Two oceans, two taxa and one mode of development: latitudinal diversity patterns of South American crabs and test for possible causal processes. *Ecol. Lett.*, 6, 420–427.
- Barneche, D.R., Rezende, E.L., Parravicini, V., Maire, E., Edgar, G.J., Stuart-Smith, R.D., *et al.* (2019). Body size, reef area and temperature predict global reef-fish species richness across spatial scales. *Glob. Ecol. Biogeogr.*, 28, 315–327.
- Bowen, B.W., Rocha, L.A., Toonen, R.J. & Karl, S.A. (2013). The origins of tropical marine biodiversity. *Trends Ecol. Evol.*, 28, 359–366.
- Bradbury, I.R., Hubert, S., Higgins, B., Borza, T., Bowman, S., Paterson, I.G., *et al.* (2010). Parallel adaptive evolution of Atlantic cod on both sides of the Atlantic Ocean in response to temperature. *Proc. R. Soc. B Biol. Sci.*, 277, 3725–3734.
- Bradbury, I.R., Laurel, B., Snelgrove, P.V., Bentzen, P. & Campana, S.E. (2008). Global patterns in marine dispersal estimates: the influence of geography, taxonomic category and life history. *Proc. R. Soc. B Biol. Sci.*, 275, 1803–1809.
- Brockington, S. & Clarke, A. (2001). The relative influence of temperature and food on the metabolism of a marine invertebrate. *J. Exp. Mar. Bio. Ecol.*, 258, 87–99.
- Chaudhary, C., Saeedi, H. & Costello, M.J. (2016). Bimodality of latitudinal gradients in marine species richness. *Trends Ecol. Evol.*, 31, 670–676.
- Chust, G., Villarino, E., Chenuil, A., Irigoien, X., Bizsel, N., Bode, A., *et al.* (2016). Dispersal similarly shapes both population genetics and community patterns in the marine realm. *Sci. Rep.*, 6, 28730.

- Clarke, A. (1992). Is there a latitudinal diversity cline in the sea? *Trends Ecol. Evol.*, 7, 286–287.
- Crame, J.A. (2001). Taxonomic diversity gradients through geological time. *Divers. Distrib.*, 7, 175–189.
- Crame, J.A., McGowan, A.J. & Bell, M.A. (2018). Differentiation of high-latitude and polar marine faunas in a greenhouse world. *Glob. Ecol. Biogeogr.*, 27, 518–537.
- Dobzhansky, T. (1950). Evolution in the tropics. *Am. Sci.*, 38, 209–221.
- Drazen, J.C. & Haedrich, R.L. (2012). A continuum of life histories in deep-sea demersal fishes. *Deep. Res. Part I Oceanogr. Res. Pap.*, 61, 34–42.
- Dynesius, M. & Jansson, R. (2000). Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proc. Natl. Acad. Sci.*, 97, 9115–9120.
- Ellingson, R.A. & Krug, P.J. (2016). Reduced genetic diversity and increased reproductive isolation follow population-level loss of larval dispersal in a marine gastropod. *Evolution*, 70, 18–37.
- Evans, K.L. & Gaston, K.J. (2005). Can the evolutionary-rates hypothesis explain species-energy relationships? *Funct. Ecol.*, 19, 899–915.
- Fernández, M., Astorga, A., Navarrete, S.A., Valdovinos, C. & Marquet, P.A. (2009). Deconstructing latitudinal species richness patterns in the ocean: does larval development hold the clue? *Ecol. Lett.*, 12, 601–611.
- Fine, P. (2015). Ecological and evolutionary drivers of geographic variation in species diversity. *Annu. Rev. Ecol. Evol. Syst.*, 46, 369–392.
- Fischer, A.G. (1960). Latitudinal variations in organic diversity. *Evolution*, 14, 64–81.
- Fraser, R.H. & Currie, D.J. (1996). The species richness-energy hypothesis in a system where historical factors are thought to prevail: coral reefs. *Am. Nat.*, 148, 138–159.
- Freestone, A.L., Osman, R.W., Ruiz, G.M. & Torchin, M.E. (2011). Stronger predation in the tropics shapes species richness patterns in marine communities. *Ecology*, 92, 983–993.
- Garcia, V.B., Lucifora, L.O. & Myers, R.A. (2008). The importance of habitat and life history to extinction risk in sharks, skates, rays and chimaeras. *Proc. R. Soc. B Biol. Sci.*, 275, 83–89.
- Gillman, L.N. & Wright, S.D. (2014). Species richness and evolutionary speed: the influence

- of temperature, water and area. *J. Biogeogr.*, 41, 39–51.
- Gillman, L.N., Wright, S.D., Cusens, J., McBride, P.D., Malhi, Y. & Whittaker, R.J. (2015). Latitude, productivity and species richness. *Glob. Ecol. Biogeogr.*, 24, 107–117.
- Goldberg, E.E., Roy, K., Lande, R. & Jablonski, D. (2005). Diversity, endemism, and age distributions in macroevolutionary sources and sinks. *Am. Nat.*, 165, 623–633.
- González-Bergonzoni, I., Meerhoff, M., Davidson, T.A., Teixeira-de Mello, F., Baattrup-Pedersen, A. & Jeppesen, E. (2012). Meta-analysis shows a consistent and strong latitudinal pattern in fish omnivory across ecosystems. *Ecosystems*, 15, 492–503.
- Grady, J. M., Maitner, B.S., Winter, A.S., Kaschner, K., Tittensor, D.P., Record, S., *et al.* (2019). Metabolic asymmetry and the global diversity of marine predators. *Science*, 363, eaat4220.
- Gray, J.S. (2002). Species richness of marine soft sediments. *Mar. Ecol. Prog. Ser.*, 244, 285–297.
- Hecht, A.D. & Agan, B. (1972). Diversity and age relationships in recent and miocene bivalves. *Syst. Biol.*, 21, 308–312.
- Heilmayer, O., Honnen, C., Jacob, U., Chiantore, M., Cattaneo-Vietti, R. & Brey, T. (2005). Temperature effects on summer growth rates in the Antarctic scallop, *Adamussium colbecki*. *Polar Biol.*, 28, 523–527.
- Hillebrand, H. (2004a). On the generality of the latitudinal diversity gradient. *Am. Nat.*, 163, 192–211.
- Hillebrand, H. (2004b). Strength, slope and variability of marine latitudinal gradients. *Mar. Ecol. Prog. Ser.*, 273, 251–267.
- Hu, Y., Ghigliotti, L., Vacchi, M., Pisano, E., Detrich, H.W. & Albertson, R.C. (2016). Evolution in an extreme environment: developmental biases and phenotypic integration in the adaptive radiation of antarctic notothenioids. *BMC Evol. Biol.*, 16, 142.
- Jablonski, D. (1986). Larval ecology and macroevolution in marine invertebrates. *Bull. Mar. Sci.*, 39, 565–587.
- Jablonski, D. (1993). The tropics as a source of evolutionary novelty through geological time. *Nature*, 364, 142–144.
- Jablonski, D., Belanger, C.L., Berke, S.K., Huang, S., Krug, A.Z., Roy, K., *et al.* (2013). Out of the tropics, but how? Fossils, bridge species, and thermal ranges in the dynamics of the marine latitudinal diversity gradient. *Proc. Natl. Acad. Sci.*, 110, 10487–10494.

- Jablonski, D., Huang, S., Roy, K. & Valentine, J.W. (2017). Shaping the latitudinal diversity gradient: new perspectives from a synthesis of paleobiology and biogeography. *Am. Nat.*, 189, 1–12.
- Jablonski, D., Kaustuv, R. & Valentine, J.W. (2006). Out of the tropics: evolutionary diversity gradient. *Science*, 314, 102–106.
- Jablonski, D. & Roy, K. (2003). Geographical range and speciation in fossil and living molluscs. *Proc. R. Soc. B Biol. Sci.*, 270, 401–406.
- Janzen, D.H. (1967). Why mountain passes are higher in the tropics. *Am. Nat.*, 101, 233–249.
- Jones, G.P., Almany, G.R., Russ, G.R., Sale, P.F., Steneck, R.S., Van Oppen, M.J.H., *et al.* (2009). Larval retention and connectivity among populations of corals and reef fishes: history, advances and challenges. *Coral Reefs*, 28, 307–325.
- Kinlock, N.L., Prowant, L., Herstoff, E.M., Foley, C.M., Akin-Fajiye, M., Bender, N., *et al.* (2018). Explaining global variation in the latitudinal diversity gradient: meta-analysis confirms known patterns and uncovers new ones. *Glob. Ecol. Biogeogr.*, 27, 125–141.
- Knudsen, R., Siwertsson, A., Adams, C.E., Garduño-Paz, M., Newton, J. & Amundsen, P.A. (2011). Temporal stability of niche use exposes sympatric Arctic charr to alternative selection pressures. *Evol. Ecol.*, 25, 589–604.
- Leis, J.M., Caselle, J.E., Bradbury, I.R., Kristiansen, T., Llopiz, J.K., Miller, M.J., *et al.* (2013). Does fish larval dispersal differ between high and low latitudes? *Proc. R. Soc. B Biol. Sci.*, 280, 20130327.
- Longo, G.O., Hay, M.E., Ferreira, C.E.L. & Floeter, S.R. (2019). Trophic interactions across 61 degrees of latitude in the Western Atlantic. *Glob. Ecol. Biogeogr.*, 28, 107–117.
- Macpherson, E. (2002). Large-scale species-richness gradients in the Atlantic Ocean. *Proc. R. Soc. B Biol. Sci.*, 269, 1715–1720.
- Marshall, D.J., Krug, P.J., Kupriyanova, E., Byrne, M. & Emlet, R.B. (2012). The biogeography of marine invertebrate life histories. *Annu. Rev. Ecol. Evol. Syst.*, 43, 97–114.
- Martin, P.R., Bonier, F. & Tewksbury, J.J. (2007). Revisiting Jablonski (1993): cladogenesis and range expansion explain latitudinal variation in taxonomic richness. *J. Evol. Biol.*, 20, 930–936.
- Matschiner, M., Hanel, R. & Salzburger, W. (2011). On the origin and trigger of the notothenioid adaptive radiation. *PLoS One*, 6, e18911.

- Menegotto, A. & Rangel, T.F. (2018). Mapping knowledge gaps in marine diversity reveals a latitudinal gradient of missing species richness. *Nat. Commun.*, 9, 4713.
- Menge, B.A., Chan, F., Dudas, S., Eerkes-Medrano, D., Grorud-Colvert, K., Heiman, K., *et al.* (2009). Terrestrial ecologists ignore aquatic literature: asymmetry in citation breadth in ecological publications and implications for generality and progress in ecology. *J. Exp. Mar. Bio. Ecol.*, 377, 93–100.
- Mittelbach, G.G., Schemske, D.W., Cornell, H. V., Allen, A.P., Brown, J.M., Bush, M.B., *et al.* (2007). Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecol. Lett.*, 10, 315–331.
- Modica, M.V., Russini, V., Fassio, G. & Oliverio, M. (2017). Do larval types affect genetic connectivity at sea? Testing hypothesis in two sibling marine gastropods with contrasting larval development. *Mar. Environ. Res.*, 127, 92–101.
- Mora, C., Tittensor, D.P. & Myers, R.A. (2008). The completeness of taxonomic inventories for describing the global diversity and distribution of marine fishes. *Proc. R. Soc. B Biol. Sci.*, 275, 149–155.
- Moss, D.K., Ivany, L.C., Judd, E.J., Cummings, P.W., Bearden, C.E., Kim, W.-J., *et al.* (2016). Lifespan, growth rate, and body size across latitude in marine Bivalvia, with implications for Phanerozoic evolution. *Proc. R. Soc. B Biol. Sci.*, 283, 867–870.
- Munch, S.B. & Salinas, S. (2009). Latitudinal variation in lifespan within species is explained by the metabolic theory of ecology. *Proc. Natl. Acad. Sci.*, 106, 13860–13864.
- Munguia, P. & Ojanguren, A.F. (2015). Bridging the gap in marine and terrestrial studies. *Ecosphere*, 6, 1–4.
- Norris, R.D. & Hull, P.M. (2012). The temporal dimension of marine speciation. *Evol. Ecol.*, 26, 393–415.
- O'Connor, M.I., Bruno, J.F., Gaines, S.D., Halpern, B.S., Lester, S.E., Kinlan, B.P., *et al.* (2007). Temperature control of larval dispersal and the implications for marine ecology, evolution, and conservation. *Proc. Natl. Acad. Sci.*, 104, 1266–71.
- O'Hara, T.D., Hugall, A.F., Woolley, S.N.C., Bribiesca-Contreras, G. & Bax, N.J. (2019). Contrasting processes drive ophiuroid phylodiversity across shallow and deep seafloors. *Nature*, 565, 636–639.
- Parmesan, C., Gaines, S., Gonzalez, L., Kaufman, D.M., Kingsolver, J., Townsend Peterson, A., *et al.* (2005). Empirical perspectives on species borders: from traditional biogeography to global change. *Oikos*, 108, 58–75.

- Peck, L.S., Massey, A., Thorne, M.A.S. & Clark, M.S. (2009). Lack of acclimation in *Ophionotus victoriae*: brittle stars are not fish. *Polar Biol.*, 32, 399–402.
- Peck, L.S., Morley, S.A., Richard, J. & Clark, M.S. (2014). Acclimation and thermal tolerance in Antarctic marine ectotherms. *J. Exp. Biol.*, 217, 16–22.
- Pianka, E.R.. (1966). Latitudinal gradients in species diversity: a review of concepts. *Am. Nat.*, 100, 33–46.
- Pörtner, H.O., Peck, L. & Somero, G. (2007). Thermal limits and adaptation in marine Antarctic ectotherms: an integrative view. *Philos. Trans. R. Soc. B Biol. Sci.*, 362, 2233–2258.
- Powell, M.G. (2007). Latitudinal diversity gradients for brachiopod genera during late Palaeozoic time: links between climate, biogeography and evolutionary rates. *Glob. Ecol. Biogeogr.*, 16, 519–528.
- Rabosky, D.L., Chang, J., Title, P.O., Cowman, P.F., Sallan, L., Friedman, M., *et al.* (2018). An inverse latitudinal gradient in speciation rate for marine fishes. *Nature*, 559, 392–395.
- Reusch, T.B.H. (2014). Climate change in the oceans: evolutionary versus phenotypically plastic responses of marine animals and plants. *Evol. Appl.*, 7, 104–122.
- Rex, M. a., Stuart, C.T., Hessler, R.R., Allen, J. a., Sanders, H.L. & Wilson, G.D.F. (1993). Global-scale latitudinal patterns of species diversity in the deep-sea benthos. *Nature*, 365, 636–639.
- Ricklefs, R.E. (2004). A comprehensive framework for global patterns in biodiversity. *Ecol. Lett.*, 7, 1–15.
- Ridgway, I.D., Richardson, C.A. & Austad, S.N. (2011). Maximum shell size, growth rate, and maturation age correlate with longevity in bivalve molluscs. *Journals Gerontol. - Ser. A Biol. Sci. Med. Sci.*, 66 A, 183–190.
- Rochowski, B.E.A., Walker, T.I. & Day, R.W. (2015). Geographical variability in life-history traits of a midslope dogfish: the brier shark *Deania calcea*. *J. Fish Biol.*, 87, 728–747.
- Rohde, K. (1978). Latitudinal gradients in species-diversity and their causes. *Biol. Zent. Bl.*, 97, 393–403.
- Rohde, K. (1992). Latitudinal gradients in species diversity : the search for the primary cause. *Oikos*, 65, 514–527.
- Rosenzweig, M.L. (1992). Species diversity gradients: we know more and less than we

- thought. *J. Mammology*, 73, 715–730.
- Roy, K. & Goldberg, E.E. (2007). Origination, extinction, and dispersal: integrative models for understanding present-day diversity gradients. *Am. Nat.*, 170 Suppl, S71–S85.
- Roy, K., Jablonski, D. & Valentine, J.W. (2000). Dissecting latitudinal diversity gradients: functional groups and clades of marine bivalves. *Proc. R. Soc. B Biol. Sci.*, 267, 293–299.
- Roy, K., Jablonski, D., Valentine, J.W., Rosenberg, G. & Osenberg, G.A.R.Y.R. (1998). Marine latitudinal diversity gradients: tests of causal hypotheses. *Proc. Natl. Acad. Sci.*, 95, 3699–3702.
- Sanders, H.L. (1968). Marine benthic diversity: a comparative study. *Am. Nat.*, 102, 243–282.
- Schemske, D. (2009). Biotic interactions and speciation in the tropics. In: *Speciation and Patterns of Diversity* (eds. Butlin, R., Bridle, J. & Schluter, D.). Cambridge University Press, pp. 219–240.
- Schemske, D.W. (2002). Ecological and evolutionary perspectives on the origins of tropical diversity. In: *Foundations of Tropical Forest Biology: Classic Papers with Commentaries* (eds. Chazdon, R.L. & Whitmore, T.C.). University of Chicago Press, pp. 163–173.
- Schemske, D.W., Mittelbach, G.G., Cornell, H. V., Sobel, J.M. & Roy, K. (2009). Is there a latitudinal gradient in the importance of biotic interactions? *Annu. Rev. Ecol. Evol. Syst.*, 40, 245–269.
- Shanks, A.L. (2009). Pelagic larval duration and dispersal distance revisited. *Biol. Bull.*, 216, 373–385.
- Siqueira, A.C., Oliveira-Santos, L.G.R., Cowman, P.F. & Floeter, S.R. (2016). Evolutionary processes underlying latitudinal differences in reef fish biodiversity. *Glob. Ecol. Biogeogr.*, 25, 1466–1476.
- Somero, G.N. (2010). The physiology of climate change: how potentials for acclimatization and genetic adaptation will determine “winners” and “losers.” *J. Exp. Biol.*, 213, 912–920.
- Steele, D.H. (1988). Latitudinal variations in body size and species diversity in marine decapod crustaceans of the continental shelf. *Int. Rev. der gesamten Hydrobiol. und Hydrogr.*, 73, 235–246.
- Steele, J.H. (1991). Can ecological theory cross the land-sea boundary? *J. Theor. Biol.*, 153, 425–436.

- Stehli, F.G., Douglas, R.G. & Kafescioglu, I.A. (1972). Models for the evolution of planktonic foraminifera. In: *Models in Paleobiology* (ed. Schopf, T.J.M.). Freeman, Cooper & Company, pp. 116–128.
- Stehli, F.G., Douglas, R.G. & Newell, N.D. (1969). Generation and maintenance of gradients in taxonomic diversity. *Science*, 164, 947–949.
- Stehli, F.G., McAlester, A.L. & Helsley, C.E. (1967). Taxonomic diversity of recent bivalves and some implications for geology. *Geol. Soc. Am. Bull.*, 78, 455–466.
- Stevens, G.C. (1989). The latitudinal gradient in geographical range: how so many species coexist in the tropics. *Am. Nat.*, 133, 240–256.
- Sunday, J.M., Bates, A.E. & Dulvy, N.K. (2011). Global analysis of thermal tolerance and latitude in ectotherms. *Proc. R. Soc. B Biol. Sci.*, 278, 1823–1830.
- Terborgh, J. (1973). On the notion of favorableness in plant ecology. *Am. Nat.*, 107, 481–501.
- Thorson, G. (1957). Bottom communities (sublittoral or shallow Shelf). In: *Treatise on marine ecology and palaeoecology* (ed. Hedgpeth, J.W.). Geological Society of America, Baltimore, pp. 461–534.
- Tittensor, D., Mora, C., Jetz, W., Lotze, H., Ricard, D., Vanden Berghe, E., *et al.* (2010). Global patterns and predictors of marine biodiversity across taxa. *Nature*, 466, 1098–1101.
- Tittensor, D.P. & Worm, B. (2016). A neutral-metabolic theory of latitudinal biodiversity. *Glob. Ecol. Biogeogr.*, 25, 630–641.
- Torchin, M.E., Miura, O. & Hechinger, R.F. (2015). Parasite species richness and intensity of interspecific interactions increase with latitude in two wide-ranging hosts. *Ecology*, 96, 3033–3042.
- Trip, E.D.L., Clements, K.D., Raubenheimer, D. & Choat, J.H. (2014). Temperature-related variation in growth rate, size, maturation and life span in a marine herbivorous fish over a latitudinal gradient. *J. Anim. Ecol.*, 83, 866–875.
- Valdovinos, C., Navarrete, S.A. & Marquet, P.A. (2003). Mollusk species diversity in the Southeastern Pacific: why are there more species towards the pole? *Ecography*, 26, 139–144.
- Valentine, J.W. & Jablonski, D. (2015). A twofold role for global energy gradients in marine biodiversity trends. *J. Biogeogr.*, 42, 997–1005.
- Webb, T.J. (2012). Marine and terrestrial ecology: unifying concepts, revealing differences.

Trends Ecol. Evol., 27, 535–541.

Wiens, J.J. & Donoghue, M.J. (2004). Historical biogeography, ecology and species richness. *Trends Ecol. Evol.*, 19, 639–644.

Willig, M.R., Kaufman, D.M. & Stevens, R.D. (2003). Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. *Annu. Rev. Ecol. Evol. Syst.*, 34, 273–309.

Winton, M. V., Natanson, L.J., Kneebone, J., Cailliet, G.M. & Ebert, D.A. (2014). Life history of *Bathyrhaja trachura* from the eastern Bering Sea, with evidence of latitudinal variation in a deep-sea skate species. *J. Mar. Biol. Assoc. United Kingdom*, 94, 411–422.

Woolley, S.N.C., Tittensor, D.P., Dunstan, P.K., Guillera-Arroita, G., Lahoz-Monfort, J.J., Wintle, B.A., *et al.* (2016). Deep-sea diversity patterns are shaped by energy availability. *Nature*, 533, 393–396.

Woolsey, E.S., Keith, S.A., Byrne, M., Schmidt-Roach, S. & Baird, A.H. (2015). Latitudinal variation in thermal tolerance thresholds of early life stages of corals. *Coral Reefs*, 34, 471–478.

Wright, S.D., Ross, H.A., Keeling, D.J., McBride, P. & Gillman, L.N. (2011). Thermal energy and the rate of genetic evolution in marine fishes. *Evol. Ecol.*, 25, 525–530.

Yasuhara, M., Tittensor, D.P., Hillebrand, H. & Worm, B. (2017). Combining marine macroecology and palaeoecology in understanding biodiversity: microfossils as a model. *Biol. Rev.*, 92, 199–215.

CAPÍTULO 3

MAPPING KNOWLEDGE GAPS IN MARINE DIVERSITY REVEALS A LATITUDINAL GRADIENT OF MISSING SPECIES RICHNESS³

³ Menegotto, A. & Rangel, T.F. (2018). Mapping knowledge gaps in marine diversity reveals a latitudinal gradient of missing species richness. *Nature Communications*, 9, 4713.

Abstract

A reliable description of any spatial pattern in species richness requires accurate knowledge about species geographical distribution. However, sampling bias may generate artefactual absences within species range and compromise our capacity to describe biodiversity patterns. Here, we analysed the spatial distribution of 35,000 marine species (varying from copepods to sharks) to identify missing occurrences (gaps) across their latitudinal range. We find a latitudinal gradient of species absence peaking near the equator, a pattern observed in both shallow and deep waters. The tropical gap in species distribution seems consequence of reduced sampling effort at low latitudes. Overall, our results suggest that spatial gaps in species distribution are the main cause of the bimodal pattern of marine diversity. Therefore, only increasing sampling effort at low latitudes will reveal if the absence of species in the tropics, and the consequent dip in species richness, are artefacts of sampling bias or a natural phenomenon.

Introduction

Large-scale spatial patterns of species richness, such as the latitudinal gradient of biodiversity, are an emergent property of the overlap of the geographic distribution of individual species (Borregaard & Rahbek 2010). Therefore, a reliable description of the spatial patterns in species richness requires accurate knowledge of the range of all species. However, the current knowledge of the geographic distribution of most species is largely incomplete (the Wallacean shortfall; Lomolino 2004; Whittaker *et al.* 2005), especially in regions of difficult and costly accessibility (Hortal *et al.* 2015). Because knowledge gaps of species distribution are not randomly distributed across space, but cluster in undersampled regions, the spatial patterns in species richness are statistically biased. Although ecologists have developed analytical techniques to estimate total species richness in a community or region (for instance using incidence of rare species; Chao & Jost 2012), these techniques usually do not reveal the identity of the potentially missing species, nor do they consider the distribution of missing species in large-scale patterns of species richness (but see occupancy models; Dorazio *et al.* 2006; MacKenzie *et al.* 2006).

Determining whether the lack of occurrence records of a species in a given locality is a true absence or a sampling artefact is one of the main challenges faced by geographical ecologists (Anderson 2003). However, if the distribution of species can be assumed to be contiguous across space at a given scale (see Supplementary Note 1; Supplementary Fig. 1), the spatial gaps in recorded presences provide an indication of the magnitude of sampling bias in a given region. For example, if a particular species is recorded at latitudes 0°, 10°, 30° and 40°, the lack of occurrence records at latitude 20° suggests an artefactual absence, especially if the sampling effort at that latitude is known to be insufficient.

By identifying the spatial gaps in the distribution of species across latitudes it is possible to map the latitudinal distribution of potential Wallacean shortfall (Fig. 1), and to

explore how and where the missing records are affecting the species richness estimate. Conversely, the absence of a species in an area of high sampling effort and inventory completeness suggests a true gap in species distribution, which is a valuable observation to study ecological processes (Pärtel *et al.* 2011). Even so, few studies have explored the spatial distribution of missing species occurrence (Anderson 2003; Bini *et al.* 2006) and, to our best knowledge, no study has yet evaluated how the patterns of species absence may affect the richness estimate at global-scale.

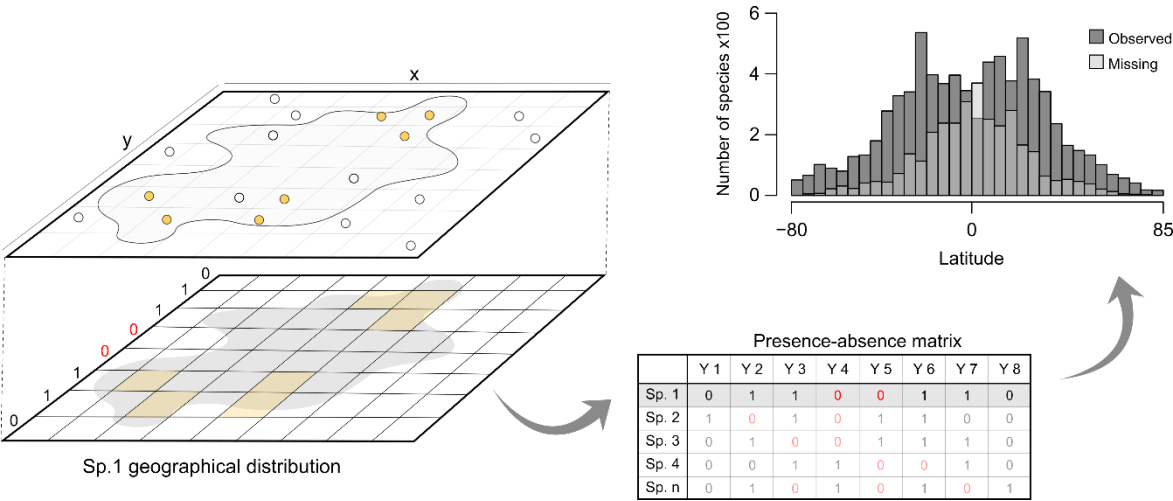


Figure 1. From species sampling to spatial richness pattern. The lack of occurrence records of a species in a location (white dots) may be caused by its true absence (geographic distribution) or by poor sampling (artefactual absence). In this example, the irregular polygon in the extruded map represents the true distribution of a species. Researchers only have access to occurrence records (yellow dots), which are then mapped in discrete grid cells (yellow cells, lower map), which may contain discontinuities in the estimated species range if the inventory is not sufficient. The observed presence (1) and absence (0) of species in latitudinal bands are recorded along the left margin of the lower map, in which red zeros indicate the spatial gaps in the latitudinal distribution of the species. By compiling occurrence data for multiple species, we can quantify not only the richness pattern but also the frequency of spatial gaps (note that colors are mixed where histograms overlap; histogram based on real Ophiuroidea data).

Here we analysed an extensive data set of more than 3 million occurrence records of 35,000 marine species, to evaluate the spatial pattern in the distributional gaps across their latitudinal range. Recent studies have proposed that the latitudinal gradient of species

richness in the marine environment is bimodal (Chaudhary *et al.* 2016; Woolley *et al.* 2016), with a dip in species richness near the equator, in marked contrast to the pattern observed in the continental realm. Here we evaluated if this dip is caused by a natural decrease in the number of species, or by an artefact of missing records of existing species. We estimated the richness of missing species by associating the observed discontinuities in species geographic ranges and the distribution of sampling effort at 5° latitudinal bands. We concluded that a bimodal pattern in marine species richness arise because of a tropical gap in species distribution, which is strongly associated with low sampling effort in the tropics.

Results

Species richness, spatial gaps and sampling efficiency

According to available records of occurrences of marine species, the latitudinal gradient in marine species richness is bimodal, with peaks at mid-latitudes (total number of species between -25° and -20°: 7,129; between 20° and 25°: 6,320) and a dip at the equator (Fig. 2a). Although not all groups have a bimodal richness distribution (copepods, stony corals and tuna fishes, have a unimodal distribution), this general pattern of species richness was observed for most taxa that we analysed (Supplementary Fig. 2).

The tropical dip in species richness coincides with a spatial gap (disjunction) in the latitudinal distribution of 11,432 species (32.80% of all species analysed). If not artefactual, such result would indicate a surprisingly common pattern of species occurring in both hemispheres, while not occurring in the tropical belt (Fig. 2a). In fact, there is a total of 5,088 species absent from the equator (more precisely between 0° and 5°), which is 1.78 times the number of species recorded at the same latitude (2,866 species). This unimodal pattern of tropical absence peaking at the equator can be observed in all taxonomic groups (Supplementary Fig. 2).

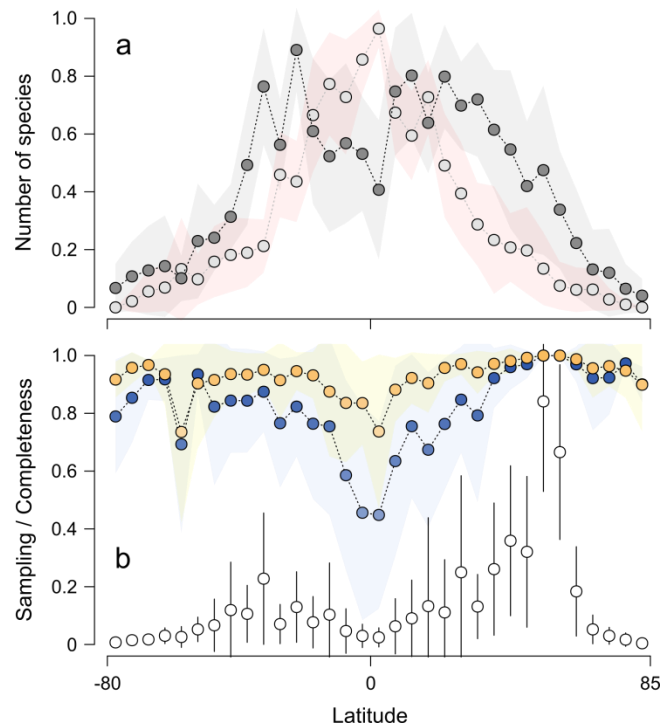


Figure 2. Patterns of latitudinal variation in species richness, spatial gaps and sampling efficiency for marine taxa. Circles represent mean values based on results of ten taxonomic groups. Shaded area and vertical bars represent standard deviation (s.d.). Data were standardized by the maximum observed value to range between 0 and 1. **(a)** Mean ($\pm$ s.d.) latitudinal variation in number of observed species (dark grey circles) and number of missing species (light grey circles). Notice that the number of missing species exceeds the number of recorded species at lower latitudes. **(b)** Mean ($\pm$ s.d.) latitudinal variation in sampling effort (number of unique sampling events; open circles) and two estimates of inventory completeness (sample coverage: orange circles; species accumulation curve: dark blue circles).

The average number of sampling events (unique combination of latitude, longitude and sampling date) was highest at mid-latitudes, especially in the northern hemisphere (Fig. 2b). In the southern mid-latitudes (between -35° and -30°), there was an average ($\pm$ standard deviation) of $9,463 \pm 15,626$ sampling events, whereas in the northern mid-latitudes (between 50° and 55°) the mean number of sampling events reached $21,644 \pm 16,455$, with a maximum of 55,346 sampling events for copepods. In fact, the number of sampling events in northern mid-latitudes was substantially higher than in any other region, for all taxonomic groups (Supplementary Fig. 3). Conversely, the mean number of sampling events at the

equator was not even higher than 700 (between -5° and 0° : 592 ± 574 ; between 0° and 5° : 628 ± 835), with a maximum of 2,688 samplings events for tuna fishes.

To evaluate if the spatial pattern of species absence is a byproduct of the unequal sampling effort we applied two different estimates of inventory completeness for each latitudinal band (sample coverage and species accumulation curve; see ‘Methods’ section). As expected, both estimates of inventory completeness revealed an unequal sampling efficiency across the globe. On average, the level of completeness was generally high at mid-latitudes, especially in those with the highest sampling effort (Fig. 2b). In the northern mid-latitudes, for example, the estimates of inventory completeness reached values near 1.0 for almost all groups analysed. Conversely, most taxa remain vastly undersampled at the equator. Considering both completeness estimation methods (sample coverage and species accumulation curve), only vertebrate fishes have an inventory completeness relatively high at the equator (above 0.9; Supplementary Figs 4-5). However, for most taxonomic groups the level of completeness was lower than 0.8 (0.74 ± 0.27) as estimated by sample coverage, and lower than 0.5 (0.45 ± 0.33) as estimated by species accumulation curve (Fig. 2b). The inventory completeness of phylum Porifera, for instance, was estimated at 0.22 by sample coverage, and approximately zero by species accumulation curve (Supplementary Figs 4-5). We also tested how the spatial resolution of analysis affects the estimated latitudinal pattern in inventory completeness. Although completeness estimates tend to increase at lower resolutions, the tropics are found to be incompletely sampled at all spatial resolutions (Supplementary Fig. 6).

Latitudinal pattern across bathymetric strata

Because the latitudinal gradient in species richness, as well as sampling intensity, may vary across depth (Webb *et al.* 2010; Woolley *et al.* 2016), we evaluated if the general

pattern described above is consistent across three different bathymetric strata: euphotic (0-200 m), bathyal (200-2,000) and abyssal (2,000-6,500). We found that latitudinal patterns in species richness, spatial gaps, and sampling efficiency in the euphotic and bathyal strata were similar to that described for the entire ocean (Fig. 3a-d), as 89.23% of all records analysed came from shallow waters (Supplementary Note 2). The average sampling effort in the deep sea is also highest at mid-latitudes of the northern hemisphere and comparatively lower in the tropics (Fig. 3f). Thus, the distribution of spatial gaps also peaked within the tropics (maximum of 361 absences for stony corals between -20° and -15° ; Fig. 3e). However, at greater depths the spatial distribution of observed species richness was highly heterogeneous among the groups (Supplementary Figs 7-16), with no clear latitudinal pattern (Fig. 3e).

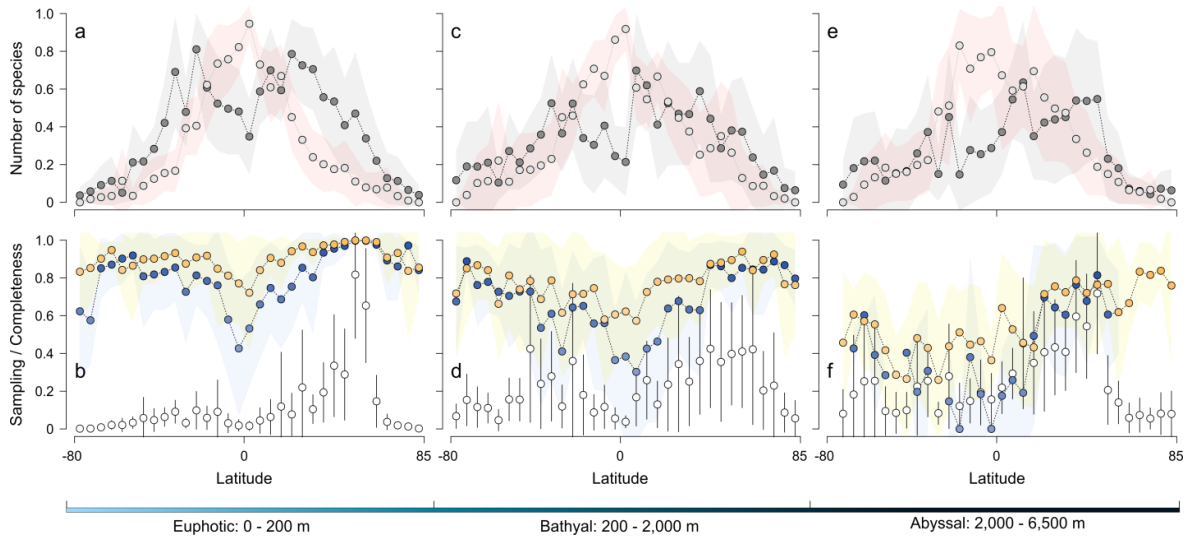


Figure 3. Patterns of latitudinal variation in species richness, spatial gaps and sampling efficiency at three different depth strata. Symbols and colors as in Fig. 2. **(a, c, e)** Mean ($\pm$ s.d.) latitudinal variation in number of observed species and number of missing species at the euphotic, bathyal and abyssal strata. **(b, d, f)** Mean ($\pm$ s.d.) latitudinal variation in sampling effort (number of unique sampling events) and two estimates of inventory completeness (sample coverage and species accumulation curve) at the euphotic, bathyal and abyssal strata. In all three depth strata the sampling effort decrease at the equator, reducing the level of completeness, number of observed species and increasing the number of spatial gaps.

Discussion

Contrary to the paradigm of a unimodal richness pattern observed in the continental realm, recent studies have proposed that the latitudinal gradient of species diversity in the marine environment is bimodal, with a dip in species richness near the equator (Chaudhary *et al.* 2016; Woolley *et al.* 2016). Indeed, by analysing an extensive data set of more than 3 million presence records, for ten different taxonomic groups, we found a bimodal pattern in species richness. However, a careful analysis of gaps in the latitudinal distribution of species revealed that the dip in species richness at lower latitudes is mostly caused by missing species records. In fact, there is an increase of gaps in species distribution toward the equator, among all ten taxonomic groups analysed, regardless of taxonomic level or body size. Although many studies have previously examined the latitudinal gradient of diversity in the marine realm (Hillebrand 2004; Tittensor *et al.* 2010; Boltovskoy & Correa 2016; Chaudhary *et al.* 2016; Woolley *et al.* 2016), our study is the first to explore the contiguity of species latitudinal distribution and to reveal a latitudinal gradient of species absences.

The absence of many marine species at the equator might suggest, at first glance, a general pattern of antitropical distribution, a phenomenon in which identical or closely related species occupy both hemispheres but are absent from the tropics (Hubbs 1952). This phenomenon is not rare in the marine system, and has been described for many taxonomic groups since the expeditions of Captain J. C. Ross in the nineteenth century (Stepanjants *et al.* 2006). It is noteworthy, however, that antitropical distribution commonly involves taxonomic groups above species level (i.e., genera or family). Although there are known marine species with antitropical distribution (Veríssimo *et al.* 2010; Havermans *et al.* 2013), there is hitherto no published account of this phenomenon as a general pattern at the species level. Indeed, if the antitropical distribution were a real biogeographical pattern, it would suggest that populations of many marine species were isolated in different hemispheres, on

course for allopatric speciation. However, the accuracy of the antitropical pattern of species distribution requires adequate sampling effort at the tropics to confirm the true absence of missing species (Ponder *et al.* 2001).

Many studies have already shown that sampling effort in the ocean is not evenly distributed across the globe (Mora *et al.* 2008; Webb *et al.* 2010; Boltovskoy & Correa 2016; Fernandez & Marques 2017). Our results support such claims and show that there are indeed fewer sampling events near the equator, even for tropical taxa (e.g., stony corals; Supplementary Fig. 3). Although sampling events are also fewer at the poles, the smaller ocean area at higher latitudes (Allen & Gillooly 2006) implies that the tropics are comparatively less sampled. Consequently, our estimates indicated a general decrease in inventory completeness toward the equator (similar to previous report for fishes; Mora *et al.* 2008). The higher sampling intensity at mid-latitudes in both hemispheres is not surprising, given the higher funding for marine research provided by developed countries at these latitudes (Mora *et al.* 2008). Conversely, most developing countries are tropical, with substantially less funding for marine biodiversity research, as indicated by the fewer number of research vessels and field stations (Costello *et al.* 2010). Thus, it seems likely that increasing sampling effort in the tropics would increase estimated tropical marine diversity by discovering new species (Costello *et al.* 2010) and by filling the tropical gaps in the distribution of marine species, therefore invalidating the bimodal pattern and the tropical dip in species richness (see Supplementary Note 3; Supplementary Fig. 2,17-18).

One of the hypotheses to explain antitropical distribution of marine species suggests that the high surface-water temperature and equatorial currents are effective barriers for species occurrence at the equator (Lindberg 1991). Therefore, this hypothesis predicts a decrease in spatial gaps (disjunct distribution) at higher ocean depths, where water is invariably colder (Tyler 2003). However, our analyses of bathymetric strata revealed that

spatial gaps in the tropics are not limited to shallow waters, but also occur in half of the deep-sea species. Thus, the apparent discontinuous distribution in abyssal species may also be the result of insufficient knowledge of the deep-sea diversity (Vinogradova 1997; Webb *et al.* 2010). In fact, our results revealed a substantial decrease in sampling events and estimates of inventory completeness at deep waters, probably reflecting the challenges of sampling those habitats (Mora *et al.* 2008). We believe that low sampling effort in tropical deep waters contributes to the artefactual absence of species in two ways: first, by reducing the chance of recording the occurrence of tropical deep-sea species at low latitudes; second, by reducing the chance of recording shallow-water polar or temperate species that could submerge to deeper waters in tropical latitudes to escape from warm temperatures (Stepanjants *et al.* 2006). In addition, our results revealed that poor sampling effort in deep waters impacts how we perceive the latitudinal gradient of marine diversity, because the observed latitudinal pattern of marine species richness is a consequence of historically higher sampling effort at shallower ocean environments.

The latitudinal pattern of spatial gaps described here is based on the most complete data repository on the distribution of marine species (Costello 2017), which has been increasingly used in analyses of marine biodiversity (Tittensor *et al.* 2010; Chaudhary *et al.* 2016). Preliminary analyses showed that adding extra data from nine other datasets on ophiuroids did not alter the general patterns described here (Supplementary Note 4; Supplementary Fig. 19; Supplementary Table 1-2). Therefore, our results show a very consistent pattern of missing species occurrence associated with low inventory completeness at the tropics. The only groups with relatively high inventory completeness and spatial gaps across the tropics are the large vertebrate fishes (i.e., elasmobranchs and tuna fishes). Although the high completeness may indicate a real absence of fish species in the tropics, it is noteworthy that sampling effort at the equator was 30 to 40 times lower than at northern

mid-latitudes for these groups (Supplementary Fig. 3). Interpolations showed that changes in sampling effort can alter considerably the tropical species richness for elasmobranchs and tuna fishes, despite the small difference in completeness estimate (Supplementary Fig. 2,17). In addition, the number of absent species is highly reduced when using information from range maps based on expert knowledge and species distribution models (Supplementary Fig. 1). Thus, the presence of spatial gaps in the tropics, associated with bimodal distribution of sampling events, may still suggest inefficiency of sampling effort for these groups. Finally, it is important to emphasize that our estimates of inventory completeness were conducted on a large spatial scale (latitudinal bands). Previous studies have shown that completeness tends to be high for large spatial scales, but decrease at finer spatial resolution (Soberón *et al.* 2007; Mora *et al.* 2008; see also Supplementary Fig. 6). Therefore, our knowledge about marine biodiversity at tropical latitudes is probably worse than reported here (Mora *et al.* 2008).

In conclusion, our results revealed that the dip in marine species diversity at the tropics is associated with a latitudinal gap in species distribution. This gap seems a consequence of low sampling at the tropics, although only the increase of sampling effort at low latitudes will confirm if the gradient in species absences is an artefact or a natural phenomenon. Improving the knowledge of marine biodiversity would require an international collaboration among research institutions of developed and developing countries (Costello *et al.* 2010), by sharing research facilities for field work (e.g., Brazil – Japan; Sumida *et al.* 2016), as well as by hosting specialists in taxonomic groups traditionally less studied (e.g., meiofauna; Fonseca *et al.* 2014). As in Schrödinger's cat paradox, in which a hypothetical cat is simultaneously dead and alive, many marine species will remain both present (artefactual absence) and absent (real absence) in the tropics until the latitudinal bias in sampling effort is reduced. In the ecological version of the famous

paradox, we will never know the real distribution of these species until we open and look inside the box of tropical oceans.

Methods

Species-occurrence data and quality control procedures

We examined 10,072,204 records of Amphipoda, Bivalvia, Copepoda, Elasmobranchii, Gastropoda, Porifera, Rhodophyta, Scleractinia and Scombridae, obtained from the Ocean Biogeographic Information System (OBIS, <http://www.iobis.org>) and of Ophiuroidea obtained from the OBIS and additional nine datasets (Atlas of Living Australia, <https://www.ala.org.au>; Global Biodiversity Information Facility, <https://www.gbif.org>; Instituto de Ciencias del Mar y Limnología - Mexico, <http://www.icmyl.unam.mx>; Instituto de Investigaciones Marinas y Costeras - Colombia, <http://www.invemar.org.co>; MNHN Paris, <http://www.mnhn.fr>; NHM London, <http://data.nhm.ac.uk>; O'Hara *et al.* (2011); Scripps Institution of Oceanography, <https://scripps.ucsd.edu>; Smithsonian NMNH, <https://naturalhistory.si.edu>). We compiled additional data for ophiuroids to evaluate the representativeness of OBIS dataset in our further analyses. Data from OBIS represented 57.81% of all valid records for ophiuroids, and preliminary analysis showed that data from OBIS is consistent with other sources ($r > 0.89$; Supplementary Note 4; Supplementary Fig. 19; Supplementary Table 1-2). Species occurrence records were obtained using the *robis* (Provoost & Bosch 2016) and *rgbif* (Chamberlain *et al.* 2017) packages implemented in the R software (R Core Team 2018) or accessing directly the portals cited above. After retrieving the data, we eliminated records without information on the geographic coordinates, coordinates equal to zero, or records located inside continents (taken into account a 20 km buffer from the coastline). Valid coordinates were rounded to two decimal places to avoid inconsistency in numerical precision among datasets, as well as to avoid double counting

sampling events. We selected only records at level of species, and eliminated records representing higher taxonomic levels. For records classified as subspecies we considered only the species name classification. We assessed the taxonomic validity of more than 36,500 species names by checking all the names at the World Register of Marine Species (WoRMS; Horton *et al.* 2017). Whenever possible, misspelling or unaccepted species names were corrected. Records with species name not available or invalid (e.g., nomen dubium, inquirendum, temporary or nudum) were excluded. Records of fossil or non-marine species were also eliminated. Finally, we checked and excluded duplicate records (specimens with exact combination of scientific name, coordinates reference and sampling date) and records at latitudes lower than -80° (Antarctic continent) or higher than 85° (Arctic ice cap). Records with sampling date unavailable were kept if the scientific name and coordinates reference represented a unique combination in the dataset. Otherwise, the record was considered duplicate and excluded. After a thorough evaluation of data quality, a total of 3,442,702 valid records, of 34,849 species, were used in our study (see details of excluded data in Supplementary Note 5; Supplementary Fig. 20; Supplementary Table 3-4; see Supplementary Fig. 21 for spatial distribution of sampling events).

Sampling intensity and inventory completeness

To quantify the latitudinal variation in sampling effort we calculated the total number of sampling events in each latitudinal band. A sampling event was defined as a unique combination of the location where a specimen was collected (latitude and longitude) and the sampling date. Additionally, we evaluated the sampling efficiency inside each latitudinal band by using two alternative methods to estimate inventory completeness: (1) we calculated the sample coverage, which is a measure of inventory completeness based on the total number of specimens recorded, and the number of rare species (singletons and doubletons;

Chao & Jost 2012). The high number of singletons and doubletons within a latitudinal band is probably an anomaly caused by undersampling (Coddington *et al.* 2009) and indicates that many other species with low abundance or restricted spatial distribution (i.e., reduced probability of detection) may exist but have not yet been recorded. (2) We used the species accumulation curves (SACs) to build a measure of inventory completeness (Tittensor *et al.* 2010; Yang *et al.* 2013). We measured the degree of curvilinearity of the last 10% of SACs for each latitudinal band to evaluate if the slope was shallow (saturation in the sampling) or steep (low completeness), as higher sampling effort in an area leads to a higher degree of curvature of its SACs. Because only the last 10% of SACs is used in the analysis, we applied this method only for latitudinal bands with more than 40 sampling events. Both methods follow the modifications of Stropp *et al.* (2016; see details in Supplementary Note 4; Supplementary Fig. 22), estimating inventory completeness between one (complete inventory) and zero (virtually inexistent inventory). Because inventory completeness may be overestimated at large spatial resolutions (Soberón *et al.* 2007; Mora *et al.* 2008), as that used here, and our goal was to compare the sampling efficiency among latitudes, we only evaluate the latitudinal variation in inventory completeness, instead of using thresholds to categorize each latitudinal band as well or poorly sampled.

Bathymetric strata

To evaluate if the latitudinal variation in species richness, spatial gaps and sampling effort are consistent across marine depth strata we partitioned all the species records in three bathymetric strata: euphotic (0-200 m), bathyal (200-2,000) and abyssal (2,000-6,500). These depth strata were selected to replicate the methodology used by Woolley *et al.* (2016). However, we used these alternative denominations, instead of continental shelf, upper slope and abyssal plane, to encompass pelagic species, for which sample depth may not necessarily

reflect the depth of the sea floor. Depth information was extracted from all records for which sample depth was recorded. When sample depth was not available for a benthic species, we used the latitude and longitude of the record to determine the depth of the sea floor at that location using ETOPO1 1 arc-minute global bathymetry (Amante & Eakins 2009). Overall, sample depth information was initially available for only 66.6% of the retrieved data. Using this additional data processing we were able to gather sample depth information for all benthic species and 90.08% of all records.

Acknowledgements

A.M. is supported by a PhD scholarship provided by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. T.F.R. has been continuously supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, grant PQ309550/2015-7). This project is also supported by INCT in Ecology, Evolution and Biodiversity Conservation, funded by MCTIC/CNPq (grant 465610/2014-5) and FAPEG (grant 201810267000023).

References

- Allen, A.P. & Gillooly, J.F. (2006). Assessing latitudinal gradients in speciation rates and biodiversity at the global scale. *Ecol. Lett.*, 9, 947–954.
- Amante, C. & Eakins, B.W. (2009). ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. NOAA Technical Memorandum NESDIS NGDC-24. National Geophysical Data Center, NOAA.
- Anderson, R.P. (2003). Real vs. artefactual absences in species distributions: tests for *Oryzomys albigularis* (Rodentia: Muridae) in Venezuela. *J. Biogeogr.*, 30, 591–605.
- Bini, L.M., Diniz-Filho, J.A.F., Rangel, T.F.L.V.B., Bastos, R.P. & Pinto, M.P. (2006). Challenging Wallacean and Linnean shortfalls: knowledge gradients and conservation planning in a biodiversity hotspot. *Divers. Distrib.*, 12, 475–482.

- Boltovskoy, D. & Correa, N. (2016). Planktonic equatorial diversity troughs: fact or artifact? Latitudinal diversity gradients in Radiolaria. *Ecology*, 98, 112–124.
- Borregaard, M.K. & Rahbek, C. (2010). Dispersion fields, diversity fields and null models: uniting range sizes and species richness. *Ecography*, 33, 402–407.
- Chamberlain, S., Barve, V. & McGlinn, D. (2017). rgbif: Interface to the Global “Biodiversity” Information Facility “API”. R package version 0.9.7. Available at: <https://CRAN.R-project.org/package=rgbif>
- Chao, A. & Jost, L. (2012). Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology*, 93, 2533–2547.
- Chaudhary, C., Saeedi, H. & Costello, M.J. (2016). Bimodality of latitudinal gradients in marine species richness. *Trends Ecol. Evol.*, 31, 670–676.
- Coddington, J.A., Agnarsson, I., Miller, J.A., Kuntner, M. & Hormiga, G. (2009). Undersampling bias: The null hypothesis for singleton species in tropical arthropod surveys. *J. Anim. Ecol.*, 78, 573–584.
- Costello, M. et al. (2017). Methods for the study of marine biodiversity. In: *The GEO handbook on biodiversity observation networks* (eds. Walters, M. & Scholes, R.J.). Springer Open, Cham, Switzerland, pp. 129–163.
- Costello, M.J., Coll, M., Danovaro, R., Halpin, P., Ojaveer, H. & Miloslavich, P. (2010). A census of marine biodiversity knowledge, resources, and future challenges. *PLoS One*, 5, e12110.
- Dorazio, R.M., Royle, J.A., Söderström, B. & Glimskär, A. (2006). Estimating species richness and accumulation by modeling species occurrence and detectability. *Ecology*, 87, 842–854.
- Fernandez, M.O. & Marques, A.C. (2017). Diversity of diversities: a response to Chaudhary, Saeedi, and Costello. *Trends Ecol. Evol.*, 32, 232–234.
- Fonseca, G., Norenburg, J. & Di Domenico, M. (2014). Editorial: diversity of marine meiofauna on the coast of Brazil. *Mar. Biodivers.*, 44, 459–462.
- Havermans, C., Sonet, G., d’Udekem d’Acoz, C., Nagy, Z.T., Martin, P., Brix, S., et al. (2013). Genetic and morphological divergences in the cosmopolitan deep-sea amphipod *Eurythenes gryllus* reveal a diverse abyss and a bipolar species. *PLoS One*, 8, e74218.
- Hillebrand, H. (2004). Strength, slope and variability of marine latitudinal gradients. *Mar. Ecol. Prog. Ser.*, 273, 251–267.

- Hortal, J., de Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M. & Ladle, R.J. (2015). Seven shortfalls that beset large-scale knowledge of biodiversity. *Annu. Rev. Ecol. Evol. Syst.*, 46, 523–549.
- Horton, T., Kroh, A., Bailly, N., Boury-Esnault, N., Brandão, S.N., Costello, M.J., *et al.* (2017). *World Register of Marine Species*. Available at: <http://www.marinespecies.org>. Last accessed 24 February 2017.
- Hubbs, C.L. (1952). Antitropical distribution of fishes and other organisms. In: *Proceedings of the Seventh Pacific Science Congress*. R. E. Owen, Government Printer, Wellington, pp. 324–329.
- Lindberg, D.R. (1991). Marine biotic interchange between the Northern and Southern Hemispheres. *Paleobiology*, 17, 308–324.
- Lomolino, M.. (2004). Conservation biogeography. In: *Frontiers of biogeography* (eds. Lomolino, V. & Heaney, L.R.). Sinauer Associates, Sunderland, pp. 293–296.
- MacKenzie, D.L., Nichols, J.D., Royle, J.D., Pollock, K.H., Bailey, L.L. & Hines, J.E. (2006). *Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence*. Elsevier.
- Mora, C., Tittensor, D.P. & Myers, R.A. (2008). The completeness of taxonomic inventories for describing the global diversity and distribution of marine fishes. *Proc. R. Soc. B Biol. Sci.*, 275, 149–155.
- O’Hara, T.D., Rowden, A.A. & Bax, N.J. (2011). A Southern Hemisphere bathyal fauna is distributed in latitudinal bands. *Curr. Biol.*, 21, 226–230.
- Pärtel, M., Szava-Kovats, R. & Zobel, M. (2011). Dark diversity: shedding light on absent species. *Trends Ecol. Evol.*, 26, 124–128.
- Ponder, W.F., Carter, G.A., Flemons, P. & Chapman, R.R. (2001). Evaluation of museum collection data for use in biodiversity assessment. *Conserv. Biol.*, 15, 648–657.
- Provoost, P. & Bosch, S. (2016). robis: R client for the OBIS API. R package version 0.1.5. Available at: <https://CRAN.R-project.org/package=robis>
- R Development Core Team (2018). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Soberón, J., Jiménez, R., Golubov, J. & Koleff, P. (2007). Assessing completeness of biodiversity databases at different spatial scales. *Ecography*, 30, 152–160.
- Stepanjants, S.D., Cortese, G., Kruglikova, S.B. & Bjørklund, K.R. (2006). A review of bipolarity concepts: history and examples from Radiolaria and Medusozoa (Cnidaria).

Mar. Biol. Res., 2, 200–241.

- Stropp, J., Ladle, R.J., Malhado, A.C.M., Hortal, J., Gaffuri, J., H. Temperley, W., *et al.* (2016). Mapping ignorance: 300 years of collecting flowering plants in Africa. *Glob. Ecol. Biogeogr.*, 25, 1085–1096.
- Sumida, P.Y.G., Alfaro-Lucas, J.M., Shimabukuro, M., Kitazato, H., Perez, J.A.A., Soares-Gomes, A., *et al.* (2016). Deep-sea whale fall fauna from the Atlantic resembles that of the Pacific Ocean. *Sci. Rep.*, 6, 22139.
- Tittensor, D., Mora, C., Jetz, W., Lotze, H., Ricard, D., Vanden Berghe, E., *et al.* (2010). Global patterns and predictors of marine biodiversity across taxa. *Nature*, 466, 1098–1101.
- Tyler, P.A. (2003). *Ecosystems of the deep oceans*. Elsevier, Amsterdam, The Netherlands.
- Veríssimo, A., McDowell, J.R. & Graves, J.E. (2010). Global population structure of the spiny dogfish *Squalus acanthias*, a temperate shark with an antitropical distribution. *Mol. Ecol.*, 19, 1651–1662.
- Vinogradova, N.G. (1997). Zoogeography of the abyssal and hadal zones. *Adv. Mar. Biol.*, 32, 325–387.
- Webb, T.J., vanden Berghe, E. & O’Dor, R. (2010). Biodiversity’s big wet secret: the global distribution of marine biological records reveals chronic under-exploration of the deep pelagic ocean. *PLoS One*, 5, e10223.
- Whittaker, R.J., Araujo, M.B., Jepson, P., Ladle, R.J., Watson, J.E.M. & Willis, K.J. (2005). Conservation biogeography: assessment and prospect. *Divers. Distrib.*, 11, 3–23.
- Woolley, S.N.C., Tittensor, D.P., Dunstan, P.K., Guillera-Arroita, G., Lahoz-Monfort, J.J., Wintle, B.A., *et al.* (2016). Deep-sea diversity patterns are shaped by energy availability. *Nature*, 533, 393–396.
- Yang, W., Ma, K. & Kreft, H. (2013). Geographical sampling bias in a large distributional database and its effects on species richness-environment models. *J. Biogeogr.*, 40, 1415–1426.

CAPÍTULO 4

SAMPLING SIMULATION IN A VIRTUAL OCEAN REVEALS STRONG SAMPLE EFFECT IN MARINE DIVERSITY PATTERNS⁴

⁴ Menegotto A., Tittensor D., Colwell R.K., Rangel T.F. Sampling simulation in a virtual ocean reveals sample effect in marine diversity patterns. *In preparation*.

Abstract

Sampling bias represents a major issue in marine macroecology. However, it is still hard to determine the extent to which this limitation is affecting our perception of marine diversity patterns. Here, we developed a computer simulation which integrates null models in macroecology with empirical information of sampling effort to evaluate experimentally the influence of the geographical sampling bias on the latitudinal gradient in marine diversity. We found that the diversity gradient has a strong imprint of sampling effort distribution. Overall, our results show that the empirical richness distribution does not seem to diverge from a null expectation after sampling simulation, and that even methods of richness estimation may fail in detect the real diversity pattern. This finding highlights the necessity of moving sampling effort to tropical latitudes. Meanwhile, we suggest that sampling bias should be considered as a null hypothesis for marine diversity patterns.

Keywords: Biodiversity gradient, computer simulation, data bias, knowledge gaps, macroecology, marine biodiversity, richness estimate, sampling bias, species distribution, Wallacean shortfall.

Introduction

The latitudinal diversity gradient is one of the most well known macroecological patterns. With few exceptions, terrestrial species richness increases from high latitudes toward the equator (Hillebrand 2004; Kinlock *et al.* 2018). Conversely, the same does not seem to hold in the marine realm. After decades of studies in marine biodiversity, the real shape of the latitudinal gradient in marine species richness is still debated. While the first descriptions of the gradient suggested a unimodal pattern, similar to that recorded on land (Thorson 1957; Fischer 1960; Stehli *et al.* 1967), other studies have repeatedly indicated that marine diversity actually peaks at mid-latitudes, composing then a bimodal pattern (Horn & Allen 1978; Clarke & Lidgard 2000; Brayard *et al.* 2005). While some studies suggest that a bimodal diversity gradient may be a general phenomenon in the marine realm (Chaudhary *et al.* 2016, 2017), some other studies suggest that such pattern may be restricted to the open ocean, as a unimodal pattern predominates in shallow coastal waters (Tittensor *et al.* 2010; Woolley *et al.* 2016).

Environmental drivers, such as productivity and temperature, have been proposed as the primary cause underlying the potential bimodal pattern in marine diversity (Brayard *et al.* 2005; Tittensor *et al.* 2010; Woolley *et al.* 2016; Chaudhary *et al.* 2017). However, the strong spatial bias in sampling effort suggests that such pattern must still be regarded with a certain skepticism (Fernandez & Marques 2017; Menegotto & Rangel 2018). Indeed, sampling bias exists since the origin of marine macroecology. For example, G. Thorson, in his seminal publication (1957), suggested that infaunal benthic communities do not exhibit a latitudinal variation in species richness. However, later, Sanders (1968) showed that such pattern was a consequence of spatial sampling bias in the tropics. Even today, in the age of global biodiversity databases, sampling bias is still a problem. Studies with fishes (Mora *et al.* 2008) and many other taxonomic groups (Menegotto & Rangel 2018) have called

attention to the reduced sampling effort in tropical and southern waters. It is true that sampling effort does not need to be evenly distributed across the globe for an adequate species diversity comparison if the sample coverage tends to be equal (Chao & Jost 2012). However, sampling effort is generally so poor in the tropics, sometimes 40 times lower than in temperate waters, that the inventory completeness is strongly reduced in such latitudes (Mora *et al.* 2008; Menegotto & Rangel 2018). Consequently, any attempt to compare latitudinal distribution of marine diversity is usually far from a robust comparison.

Studies exploring global patterns in marine diversity have employed a variety of methodological techniques to estimate species richness and control for sampling bias. Such techniques usually involve sample-based rarefaction (e.g., interpolation; Chaudhary *et al.* 2017), extrapolations based on incidence of rare species (e.g., Chao 2 estimator; Tittensor *et al.* 2010) or species distribution models (e.g., occupation models; Woolley *et al.* 2016). Using those methods some studies have found a higher diversity in mid-latitudes, which reinforces the bimodality hypothesis (Tittensor *et al.* 2010; Woolley *et al.* 2016; Chaudhary *et al.* 2017). However, Menegotto & Rangel (2018) recently showed that the currently documented tropical dip in marine species diversity is contingent upon most marine species having disjunct distributions at lower latitudes (i.e., a tropical “gap” in their ranges), as most marine species have been recorded in both hemispheres, but not in the tropics. If, however, the tropical gap is assumed to be artefactually caused by poor sampling at low latitudes, and species range are therefore made contiguous (i.e., by “filling” the tropical gap in their latitudinal ranges), a unimodal pattern in species richness emerges with peak at the equator (Menegotto & Rangel 2018). Therefore, the final verdict on the shape of latitudinal diversity gradient will only come with additional sampling effort at the tropics, which is unlikely to happen anytime soon, given the currently low investment and lacking infrastructure for marine biodiversity research in tropical countries (Costello *et al.* 2010).

Simulating the existing sampling effort in a virtual ocean, in which the distribution of virtual species is known with certainty, allows the evaluation of the impact of sampling biases on the perceived tropical gaps in species distribution, as well as its influence on the emerging latitudinal gradient in marine biodiversity. Computer simulations have been used to generate null models to test macroecological hypothesis (Colwell & Hurtt 1994; Jetz & Rahbek 2001; Carranza *et al.* 2008) and also to replicate empirical biogeographical patterns (Brayard *et al.* 2005; Rangel *et al.* 2007, 2018; Tittensor & Worm 2016; Coelho *et al.* 2018). Conceptually, these simulation models can be viewed as replacement of empirical experiments that would not be feasible owing to logistic, ethical or budget constraints (Peck 2004). Thus, under different hypothetical scenarios of species distribution in a virtual ocean, historical marine sampling may be simulated to evaluate the impact of sampling biases on the perceived pattern of marine diversity. In addition, because the true diversity of the artificial *in silico* world is known with full certainty, it is also possible to evaluate the accuracy of the statistical methods frequently used by marine macroecologists to overcome the effect of sampling bias.

Here we developed a simulation model to understand how the historical distribution of sampling events may affect our perception of the latitudinal biodiversity gradient observed in the marine realm. Our computer simulation implements two common null models in macroecology used to generate distribution of virtual species, the mid-domain and the area effect. While the mid-domain model assumes range cohesion, the area model allows spatial gaps in the latitudinal distribution of simulated species. Then, we sampled the virtual species using the empirical distribution of marine sampling effort, and compared the estimated latitudinal diversity gradient with the reference pattern (i.e., the null model), considering the effects of sampling bias and non-cohesiveness of species ranges. For each sampling simulation we also calculated the latitudinal sample coverage and the predicted

number of species at the cell scale, to evaluate the efficiency of methods used to identify and reduce spatial bias effect. Our results show that sampling bias leaves a strong imprint in marine diversity patterns, and that even richness estimation methods may not be able to indicate if the perceived tropical dip in marine diversity is a sampling artefact or a true biological phenomenon caused by species with disjunct distribution.

Methods

Model overview

Our model simulates the sampling process of marine species on a two-dimensional geographical domain, which is represented here as a gridded map of the Atlantic Ocean divided into 396 cells of $5^\circ \times 5^\circ$. Conceptually and computationally, this domain is delimited longitudinally by continents and latitudinally by polar oceans. A simulation cycle includes three steps: distributing virtual species in the domain, which produces an incidence matrix with the “reference” diversity; simulating the sampling of species, which produces an incidence matrix with the “observed” diversity; and data analyses, which is a set of estimated metrics (e.g., number of missing species and inventory completeness) based on the simulated sampling of the virtual species. Species distribution and sampling process are completely stochastic, with no link to environmental niche or detection probability. The only parameters derived from empirical data are (1) the total number of species, (2) the latitudinal distribution of sampling effort and (3) the average number of species record by sampling events. We used here ophiuroids records to estimate these three parameters.

Empirical data

Ophiuroidea has a clear bimodal pattern in species diversity (Chaudhary *et al.* 2016; Woolley *et al.* 2016; Menegotto & Rangel 2018), while many species exhibit strong evidence

of spatial gaps in their tropical distribution (Menegotto & Rangel 2018). Global occurrence records of ophiuroids were retrieved in February 2017 from different sources worldwide (Ocean Biogeographic Information System, <http://www.iobis.org>; Atlas of Living Australia, <https://www.ala.org.au>; Global Biodiversity Information Facility, <https://www.gbif.org>; Instituto de Ciencias del Mar y Limnología, <http://www.icmyl.unam.mx>; Instituto de Investigaciones Marinas y Costeras, <http://www.invemar.org.co>; MNHN Paris, <http://www.mnhn.fr>; NHM London, <http://data.nhm.ac.uk>; O'Hara *et al.* (2011); Scripps Institution of Oceanography, <https://scripps.ucsd.edu>; Smithsonian NMNH, <https://naturalhistory.si.edu>). After a thorough data cleaning process (see details in Menegotto & Rangel 2018), we selected only those records within the Atlantic Ocean, our geographical domain, in a total of 69,609 records of 624 species. The number of sampling events per 5° latitudinal band ranges from 25 in the tropics to 11,705 at mid-northern latitude, with an average of 2 species recorded at each sampling event. These values were used to configure the parameters of the simulation (see below). To illustrate the versatility of our sampling simulation we also repeated the analyses using input parameters from two additional taxa, with widely different diversity within the Atlantic Ocean, the tuna fishes (30 spp.) and gastropods (4,388 spp.).

Model dynamics

The simulation starts allocating species ranges in the geographic domain. For each species, the range is generated by expanding from a randomly chosen seed cell until a previously defined range size is reached. We drew range sizes at random from a beta distribution ($\alpha=3$ and $\beta=9$) to guarantee that the range size frequency distribution of simulated species conforms to the lognormal shape usually observed in nature (Gaston 1996). In addition, the geographic structure of simulated ranges varies according to two

different models (Jetz & Rahbek 2001). First, the mid-domain model produces the range of a species by progressively expanding from a seed cell into adjacent and unoccupied cells, until the final range size is reached. Alternatively, the area model produces ranges by randomly occupying any of the unoccupied cells of the domain, similarly to an unlimited dispersal process. Regardless of the model, the number of overlapping simulated ranges in each cell, or latitudinal band, represents the “reference” diversity.

Once the distribution of all species is generated according to the mid-domain or area model, we then simulate the historical sampling effort to generate an “observed” latitudinal diversity gradient. In our simulation, the number of sampling events at each latitudinal band is defined according to historical sampling effort at that latitude for a given target group of species (e.g., Ophiuroidea). In empirical databases, sampling events are unique combination of location (i.e., geographical coordinates) and date, which means that the same cell may receive different sampling events over time. In addition, a single sampling event does not record all species present in the entire grid cell, but only a fraction of the cell’s biota. However, as the sampling effort within a cell increases, the biological inventory of the cell approaches the completeness. Based on the historical sampling effort of Ophiuroidea, we estimated that 2 is the average number of species recorded by a single sampling event. Thus, in our simulation of sampling virtual species, for each sampling event within a latitudinal band, a grid cell is chosen at random, and 2 species are randomly chosen and recorded as sampled.

After simulating the sampling process, we estimated observed diversity (i.e., the diversity resulting from sampling record), spatial gaps in the latitudinal range of species (Menegotto & Rangel 2018) and sample coverage (Chao & Jost 2012) for all latitudes. Recent studies in marine macroecology have commonly used extrapolation of species richness on a cell-by-cell basis to overcome the sampling bias problem (Tittensor *et al.* 2010;

Woolley *et al.* 2016). Therefore, we also calculated the predicted number of species for each cell using the nonparametric Chao 2 estimator (Chao 1984) to compare the predicted richness with the reference diversity.

Finally, to estimate the minimum unbiased sampling effort necessary to reach a good taxonomic description of simulated species, and an accurate latitudinal diversity distribution, we implemented a second sampling scenario in which sampling events are uniformly distributed across all latitudinal bands, rather than replicating the biased geographical distribution of historical sampling events. We started with 25 sampling events at each latitudinal band, increasing progressively until 5000 sampling events per band, comprising a total of 200 sampling scenarios. These limits represent, respectively, the number of sampling events at the latitudinal band with the least and half of the maximum sampling effort for Ophiuroidea. Besides the observed diversity for each latitudinal band, we also calculated the spatial gaps and sample coverage, which were progressively updated as the sampling effort increased.

Data Analysis

Preliminary analyses indicated that 2,000 model runs are sufficient to reach stable average and standard deviation of observed diversity patterns (Fig. S1). Conservatively, we ran each model 10,000 times. We used the Pearson's correlation coefficient (r) to compare the reference (simulated) and observed (sampled from reference) diversity gradients, as well as empirical and virtual patterns in observed species richness, spatial gaps and sample coverage. We also calculated the correlation coefficient of the reference and observed diversity with the predicted species richness at the cell scale. All statistical analyses were conducted in the R software (R Core Team 2018).

Results

Mid-domain model

Our results show that despite the clear tropical peak in species diversity created by the mid-domain model, the sampling simulation based on empirical distribution of historical sampling effort resulted in a completely biased estimate of latitudinal diversity gradient. Overall, the observed species richness resulting from sampling simulation peaked at mid latitudes, especially in the northern hemisphere, with a dip at the equator (Fig. 1b). The sampling simulation of the virtual ocean using the historical sampling effort did not generate an accurate estimate of the reference species richness ($r = -0.04$). Ironically, it did reproduce very well the empirical diversity pattern of ophiuroids ($r = 0.89$; Fig. 1). The estimate of inventory completeness, calculated for the sampling simulation, also reproduced adequately the sample coverage estimated from empirical data ($r = 0.82$). Overall, the completeness estimates at poorly sampled latitudes were lower for the simulated dataset than for the real-world dataset. However, estimate of inventory completeness accurately identified latitudinal bands where most of the virtual species were not recorded ($r = 0.99$). As for the real world, these latitudes are located near the equator and in the southern hemisphere (Fig. 2a). Adjusting the simulation parameters to datasets of tuna fishes and gastropods yielded the same correspondence between empirical and virtual estimates (Fig. S2).

Latitudes with smaller sampling effort and low inventory completeness were strongly associated with a high frequency of spatial gaps in species range ($r = -0.92$; $r = -0.99$). Curiously, the distribution of spatial gaps resulting from sampling simulation was only moderately correlated with that observed in empirical data ($r = 0.41$), probably because of the high number of virtual species missing in southern mid-latitudes (Fig. 2b-c). However, we noted that our biogeographical simulation produced a higher species richness in high latitudes compared to empirical data (Fig. 1a-b, 2b-c). That is, the empirical diversity

distribution tends to be leptokurtic while in our simulation it was platykurtic. To evaluate if the excess of kurtosis could increase the similarity between virtual and empirical spatial gaps, we repeated the simulation increasing the probability of species originating in the tropics (i.e., the center of the domain) with a beta distribution of $\alpha = \beta = 3$ and $\alpha = \beta = 9$. Our results show that as tropical origination rate increases, the distribution of spatial gaps fits the empirical data better (Beta(3, 3): $r = 0.61$; Beta(9, 9): $r = 0.74$; Fig S3).

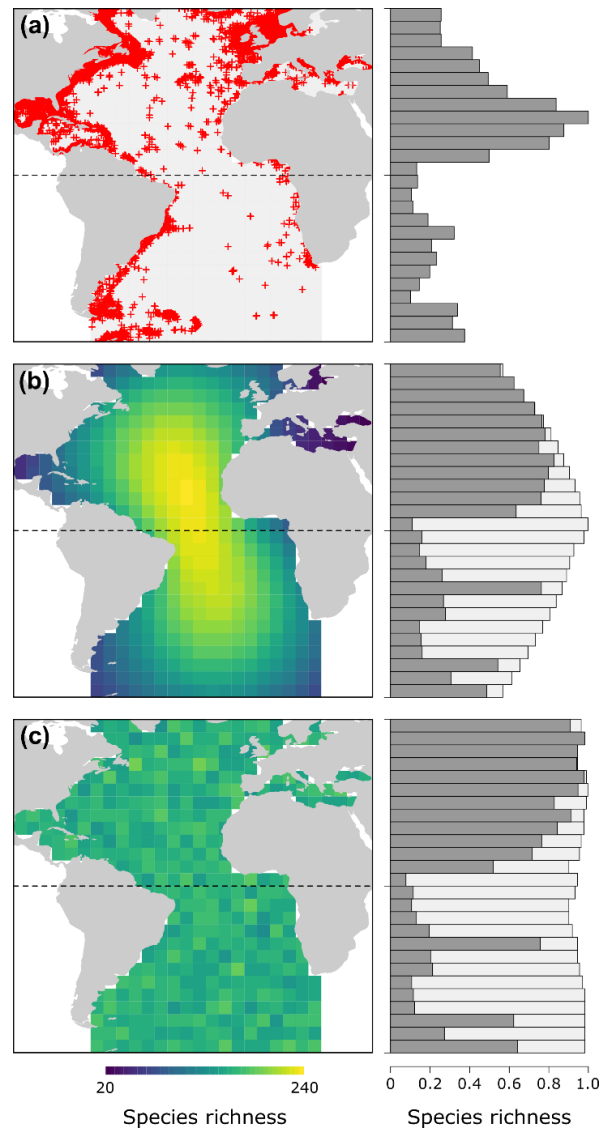


Figure 1. Sampling simulation and virtual species diversity in the Atlantic Ocean. (a) Map showing the distribution of sampling events for Ophiuroidea (red crosses) and the observed species richness per latitude. Examples of the mid-domain (b) and area model (c) show the average number of species per latitude (light grey) and the number of species recorded in the sampling simulation (dark grey). Histogram data were standardized by the maximum species richness.

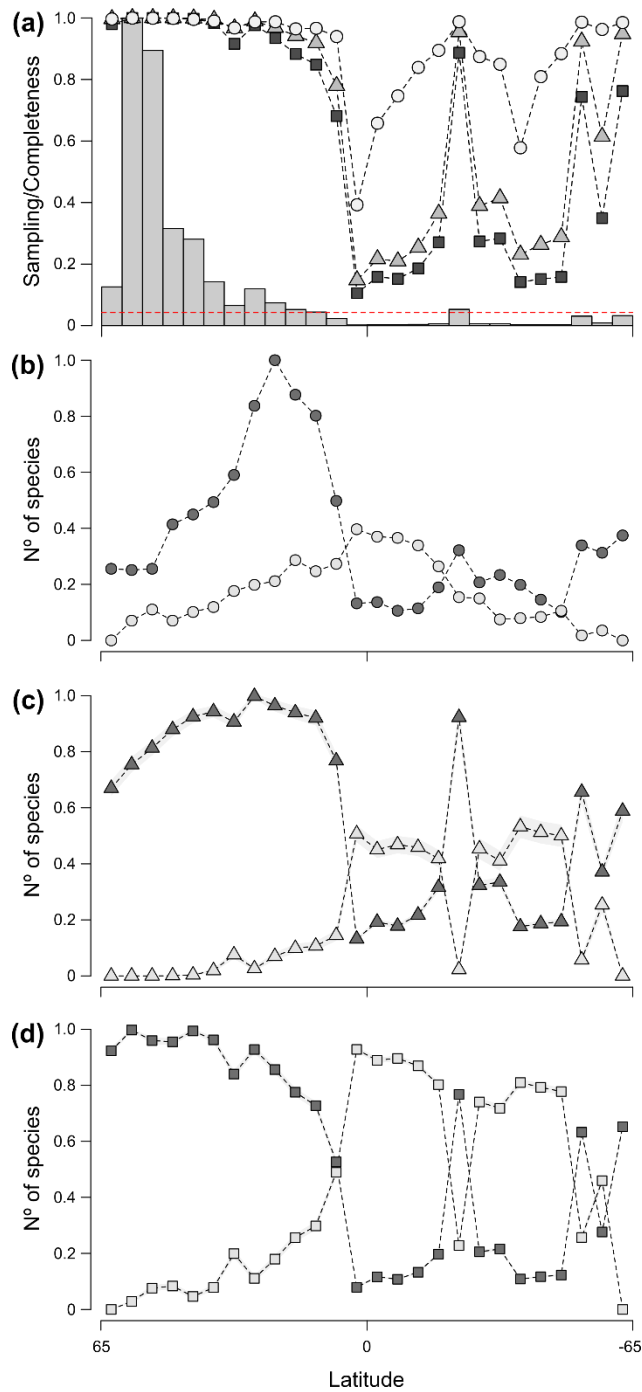


Figure 2. Inventory completeness and spatial gaps in empirical and simulated sampling records. (a) Inventory completeness estimated for the empirical data (light grey circles), mid-domain model (grey triangles) and area model (dark grey squares). Histogram represent standardized sampling effort. For comparison purposes, we indicate the mark of 500 sampling events with the dotted red line. Plots below indicate the estimated diversity (dark grey) and number of missing species (light grey) in the empirical data (b), mid-domain model (c), and area model (d). Grey shadows represent standard deviation.

The sampling simulation also did not capture well the reference diversity at the cell scale ($r = -0.01$). Therefore, we calculated the predicted species richness to evaluate if we could get close to the true diversity at each cell and reproduce accurately the latitudinal diversity gradient. However, the predicted species richness still reproduced better the observed (i.e., sampled) than the reference diversity pattern ($r = 0.80$; $r = 0.39$). An inspection of the correlation plot between the reference and predicted diversity revealed that, overall, the species richness estimator performed well. However, the estimation performed poorly at those latitudes with fewer sampling events (Fig. 3). Overall, species richness was underestimated in those strongly under-sampled cells, such as at the equator (<150 sampling events); and overestimated in those moderately under-sampled cells, as other tropical latitudes ($150 < \text{sampling events} < 400$; Fig. 4).

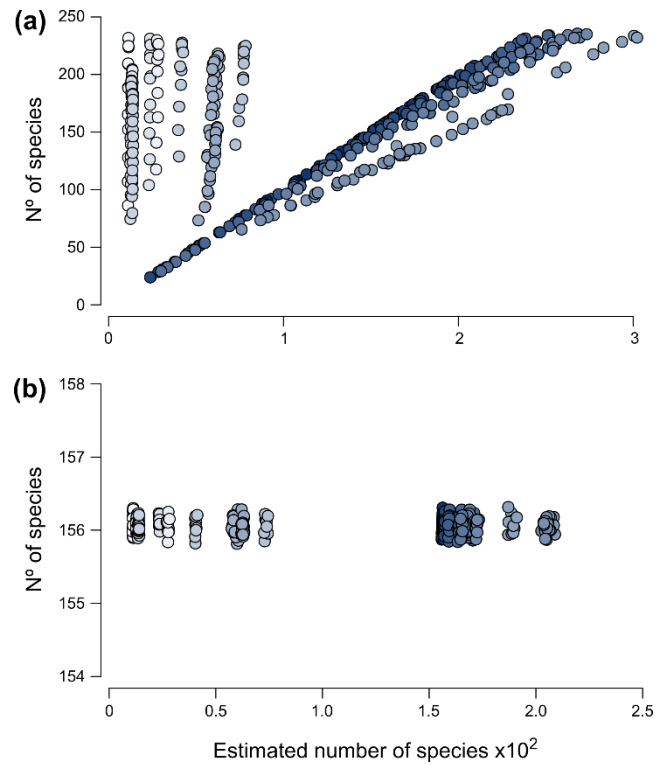


Figure 3. Accuracy of the species richness prediction. Relationship between reference and predicted diversity at cell scale for (a) the mid-domain model and (b) the area model. Color gradient indicates a variation from small (light) to high (dark) sampling effort.

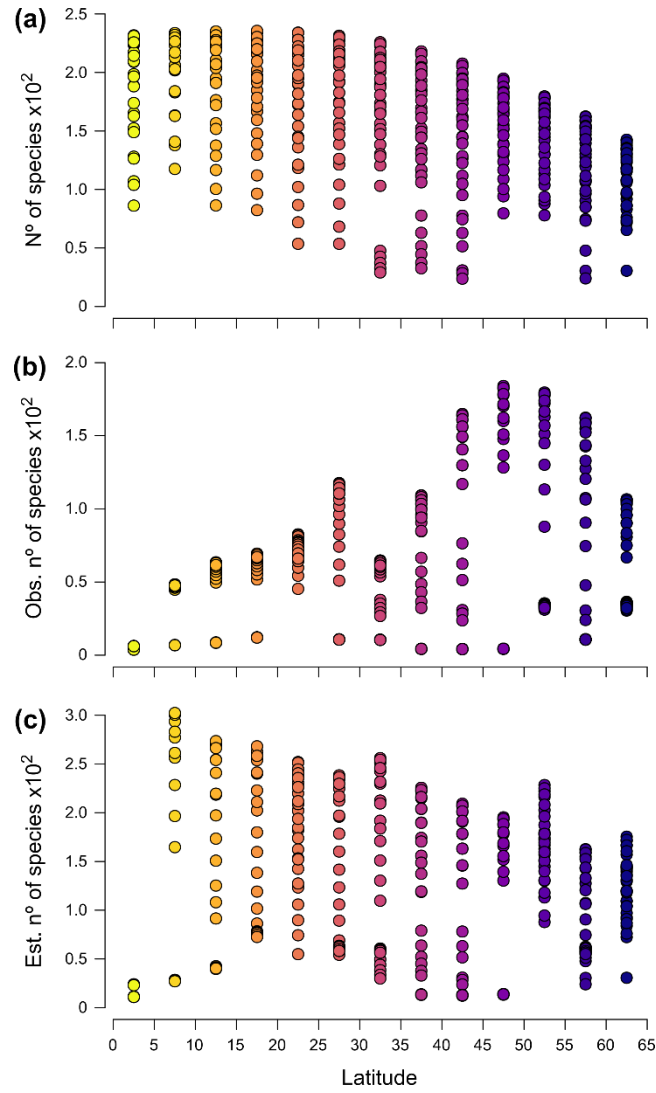


Figure 4. Latitudinal species diversity by cell in the mid-domain model. (a) Number of species in the null model (reference diversity). (b) Number of species recorded in the sampling process (observed diversity). (c) Predicted number of species (estimated diversity). Latitudes are shown in absolute values. Colors indicate variation from high (cold) to low (hot) latitudes.

Area model

In a scenario where species ranges are scattered randomly across the geographic domain, the average latitudinal diversity pattern results in a relatively flat distribution (Fig. 1c). Although the reference diversity distribution in this scenario is quite different from that created by the mid-domain effect ($r = -0.54$), there was no significant difference between the observed species richness of the two simulations ($r = 0.94$; Fig. 1b-c). Similarly, the resulting

completeness estimate and spatial gaps distribution of the two models were highly correlated (both with $r = 0.98$; Fig. 2).

Analysis of species richness by cell also confirmed that, despite the clear difference between the two simulated scenarios ($r = 0.03$), the observed diversities were highly correlated ($r = 0.98$). As in the mid-domain effect scenario, the sampling bias was the main driver of the poorly species richness prediction at the cell scale (Fig. 3). Consequently, we also detected a spuriously high correlation between the mid-domain and area model in the predicted diversity ($r = 0.79$).

Increasing sampling effort

In our virtual experiment of increasing sampling effort, few sampling events in the mid-domain effect scenario resulted in a modest diversity gradient, though the tropical peak in species diversity accurately stood out as the sampling effort increased (Fig. S4). With only 500 sampling events it was possible to record 90% of the virtual species at high latitudes. However, two times more sampling effort was needed in tropical latitudes to obtain the same sample coverage (Fig. 5a). Finally, a little more than 4,000 sampling events were necessary to record all species at each latitude (Fig. S4). Running the same experiment for the area model resulted in a species accumulation curve quite undistinguishable among the latitudes (Fig. 5b). Under this scenario, all latitudinal bands needed nearly 1,000 sampling events to record 90% of the species.

The frequency of spatial gaps in species distribution peaked under low sampling effort, reaching up to 128 species in a latitude in the mid-domain model and 398 in the area model (Fig. 6). In contrast, in our empirical database we recorded a maximum of 90 missing species in only one latitude. The number of gaps in species latitudinal range reduced exponentially as sampling effort increased. With 1,000 sampling events, which generates a

sample coverage of at least 90% for all latitudes, only 11 ± 5 (mean $\pm$ sd) species remained artefactually absent in the mid-domain model. Under the area model, where species may be naturally absent from some latitudes, the frequency of spatial gaps was reduced to 98 ± 16 species. In both scenarios, the frequency of spatial gaps was usually high in tropical latitudes, even with equal sampling effort (Fig. 6). Analysis of the proportion of missing species revealed, though, that only when sampling effort is extremely small there is more spatial gaps than species recorded (Fig. 6), as observed in empirical data (Fig. 2b). Overall, the spatial gaps/species record ratio is reduced if under-sampling is not so drastic, even when species may have latitudinal ranges naturally disjunct.

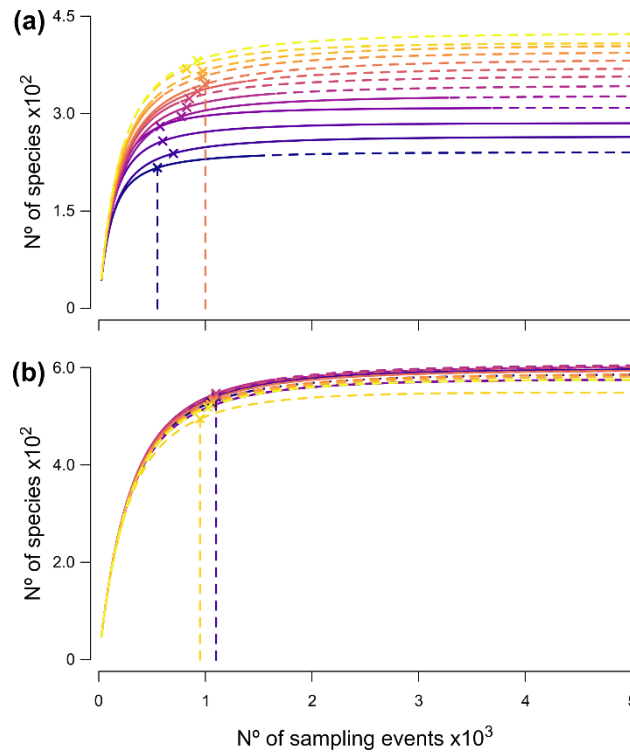


Figure 5. Species accumulation curve for 5,000 sampling events. (a) Mid-domain model. (b) Area model. Colors indicate variation from high (cold) to low (hot) latitudes. Dotted lines represent the species accumulation curve within the simulation and we indicate with the solid line the extension of real sampling effort. Crosses indicate the point where 90% of the species were recorded and vertical dotted lines highlight the minimal and maximal position. For simplicity, we show here only the results for the northern hemisphere.

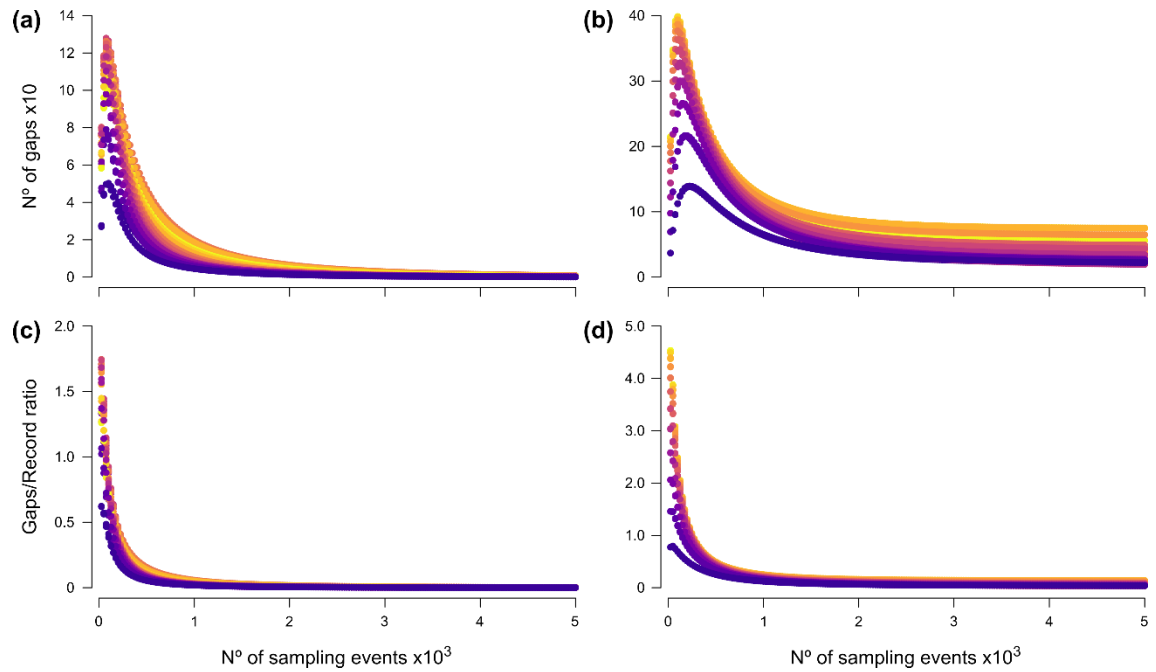


Figure 6. Changes in spatial gaps with increasing sampling effort. Number of spatial gaps for each latitude in the (a) mid-domain model and (b) area model. Plots below represent the proportion of spatial gaps by the number of species recorded in the sampled events of the (c) mid-domain model and (d) area model. Colors indicate variation from high (cold) to low (hot) latitudes.

Discussion

Sampling bias represents a major issue in marine macroecology, but it is still hard to determine the extent to which this limitation is affecting our perception of marine diversity patterns. Because the sampling effort in the tropics has been extremely poor, inferences on the true pattern of species diversity are mostly extrapolations. By assessing the diversity pattern resulting from a virtual sampling simulation, which attempts to emulate our limitation of working with observed richness data in real world, here we show how difficult it is to detect the underlying gradient in marine species richness with the currently available sample coverage.

Here we implemented a simulation model that is completely deprived of biological processes and is not designed to replicate any pattern of real-world species richness. Surprisingly, the diversity pattern resulting from our sampling simulation was more similar

to the historical record of real-world organisms (i.e., the pattern in species richness estimated empirically) than to the reference pattern of species diversity that was simulated (i.e., the true distribution of virtual species). Regardless if the original trend was unimodal or flat, when replicating the historical sampling effort the observed species richness always peaked at mid-latitudes of the northern hemisphere. The same correspondence between simulated and empirical observed diversity occurred when we set the simulation parameters with data from taxonomic groups with different number of species and sampling effort. Thus, our study does not only reveal that the simulation model succeeded in reproducing the imprint of the sampling process, but it also suggests that it is virtually impossible to use the historical sampling effort to infer the true pattern of real-world marine diversity.

Besides the richness pattern, our sampling simulation also captured well the latitudinal pattern in inventory completeness, which tends to peak at high latitudes and decrease toward the equator for many marine taxa (Mora *et al.* 2008; Menegotto & Rangel 2018). At poorly sampled latitudes the completeness estimated in the simulation was considerably lower than that estimated for the historical sampling of the real world. However, it is noteworthy that in our simulation each cell within a latitudinal band has equal chance to be sampled, whilst in the real world most of the sampling events may be restricted to few cells within a latitudinal band (e.g., due to research facilities and national funding; Costello *et al.* 2010; Meyer *et al.* 2015; see also Fig. S5). Although such process will certainly result in more information about alpha diversity, it seems to inflate the latitudinal completeness estimation and allow the rising of spatial gaps in species distribution (Fig. S6). If true, this result would indicate that the estimated tropical inventory completeness, although very low, may still be overestimated, with many more species remaining to be recorded in the tropics.

The distribution of spatial gaps resulting from the sampling simulation was not so similar to the empirical data compared to other metrics investigated here. Such difference emerged simply from the kurtosis of the latitudinal diversity curve generated by the mid-domain and area effect. Once we increased the species origination rate in the tropics, which may be a reality for marine invertebrates (Jablonski 1993; Allen & Gillooly 2006; Jablonski *et al.* 2006), the overall number of species in high latitudes approached the real curve and the correlation with the empirical distribution of spatial gaps increased. That is, the number of missing species was reduced in southern temperate waters and continued high in the tropics, fitting the empirical data. This result illustrates how the missing species gradient should emerge from an under-inventoried high tropical diversity if marine species range were indeed cohesive (Menegotto & Rangel 2018). Interestingly, our results also revealed that by just filling the recorded tropical gap in species distribution it was not possible to replicate the true species richness (Fig. S7). Therefore, assuming a high tropical diversity, sampling bias would be not only increasing the Wallacean shortfall, but also the Linnean shortfall, since many tropical species, especially the endemic ones, should remain largely unknown by science in these poorly sampled latitudes (Whittaker *et al.* 2005; Hortal *et al.* 2015).

The high correlation between spatial gaps resulting from the two simulated scenarios is curious because only under the area effect species may have natural gaps in their geographical distribution. In the mid-domain scenario such pattern only emerges from inadequate sampling effort. Even so, for both scenarios the frequency of spatial gaps was concentrated in latitudes with fewer sampling events. Although the presence of naturally disjunct distribution may be slightly higher in the tropics due to geometric constraints imposed by the shape of the domain, the poor sampling effort seems to intensify this pattern, increasing the number of missing species to a level higher than the number of recorded

species. The same proportion is observed in empirical data from lower latitudes (Menegotto & Rangel 2018). According to the results of the increasing sampling effort simulation, such pattern can only indicate an effect of sampling bias, unless, of course, that the antitropical distribution is a common feature for most of the marine species.

Our results show that in a scenario of higher tropical richness, the tropics would require twice the sampling effort of mid-latitudes to reach the same inventory completeness (c.1,000 samples in the tropics against c.500 in temperate waters for a 0.9 sample coverage), as under equal sampling effort species are more likely to be missed where species richness is higher (Colwell & Hurtt 1994). In agreement with our simulation results, analyses of empirical data have also indicated that tropical latitudes would need more samples to achieve a comparable sample coverage (Menegotto & Rangel 2018), suggesting that sampling effort should be concentrated in tropical latitudes to accurately describe the true natural latitudinal diversity pattern. Conversely, the actual distribution of sampling effort is much higher in temperate waters (maximum of 11,000 sampling events), with many tropical latitudes not reaching even 500 sampling events (minimum of 25 sampling events). It is likely that existing records from lower latitudes have not been integrated into our analyses due to gaps in digital accessible information (Meyer *et al.* 2015; Fernandez & Marques 2017), what could increase tropical sample coverage and, potentially, number of species. However, while such data is still not available, marine diversity estimates continue strongly dependent on statistical techniques to deal with sampling biases (Tittensor *et al.* 2010; Woolley *et al.* 2016; Chaudhary *et al.* 2017).

One strategy usually employed in the literature to deal with sampling bias is to use interpolation or extrapolation methods to estimate species richness (Colwell & Coddington 1994; Chao & Jost 2012). Interpolation methods are characterized by the down-sample of a heterogeneous sampling effort to an equally small sample size (but see Chao & Jost 2012

suggestion for standardization by sample coverage), and may be compared here to the first steps of our increasing sampling experiment. Although our sample coverage estimates have already shown that more samples are needed in the tropics to achieve a comparable completeness (see also Menegotto & Rangel 2018), the equal sampling size standardization to 50 samples by latitude seems to show, at first glance, a satisfactory result to detect the shape of the original diversity pattern (Fig. S6). If we are correct, this result could indicate that a bimodal pattern detected using this method is, therefore, consistent (Chaudhary *et al.* 2017). However, in our simulation each cell can be equally sampled for all latitudes, equalizing the probability of different species being detected. Analyses of empirical data, on the other hand, show that the number of sampled cells is much higher in high latitudes than in the tropics (Chaudhary *et al.* 2017; Menegotto & Rangel 2018; see also Fig. S5). This means that a higher spatial extent, which probably includes the geographical ranges of many different species, will be available for re-sampling in high latitudes, reducing inevitably the estimated species richness near the equator (see a simple experiment about that in the Fig. S6). Therefore, we suggest that interpolation methods should be used with caution to evaluate the shape of marine diversity gradients, not only respecting the sample coverage of the cells but also the size of the area compared across latitudes.

Another method used to estimate marine diversity is to extrapolate the cell-scale species richness. However, even these non-parametric richness estimators are not immune to sampling bias and may still fail to provide reliable estimates at low sampling effort (Chiarucci *et al.* 2003; Lopez *et al.* 2012). Our results reveal that the interaction between sample size and estimate deviation causes an underestimation of species richness in cells at the equator, while in adjacent tropical latitudes it could be overestimated, therefore increasing the perception of an equatorial dip in marine diversity. Because species richness estimates at poorly sampled cells are not reliable, we conjecture that the oceanic zone, where

sampling effort is even scarcer (Webb *et al.* 2010; Menegotto & Rangel 2018), may have the strongest bias. Indeed, a previous study using this method have already shown that the bimodal pattern tends to be more common in oceanic than in coastal waters (Tittensor *et al.* 2010). Yet, it is important remember that a similar result of bimodality for open ocean was also found for deep-sea species using a different approach. Woolley *et al.* (2016) detected a bimodal pattern using methods of species distribution models that account for their detectability (i.e., multi-species occupancy-detection models). Future studies should then evaluate how such models respond to the massive sampling bias in the marine realm. If proved efficient, it could reinforce the bimodality hypothesis in the oceanic zone.

Overall, methods of richness estimation showed that they are accurate if their limitations are respected. However, further studies evaluating the efficiency of these methods in more realistic scenarios are still necessary. It is known that marine species have different degrees of detectability due to physical, behavior, or life-history characteristics (MacNeil *et al.* 2008; Bates *et al.* 2015; Karenyi *et al.* 2018), and that even large species may be overlooked by traditional sampling methods (Boussarie *et al.* 2018). Conversely, in our simulation species had equal chance of detection. Although the diversity pattern observed after sampling simulation was very similar to that of the real world, a realistic scenario may be more appropriated to evaluate the complexity of species richness estimation. By acknowledging the limitation of the available data and methods we may improve them and advance our knowledge about marine diversity.

Finally, it is important to emphasize that our simulation was not designed to produce a realistic pattern of marine diversity (e.g.; Brayard *et al.* 2005), neither does it prove that marine species richness should peak at the equator. Although tropical diversity is most likely higher than observed so far, it is possible that diversity at mid-latitudes is higher than in the tropics (Brayard *et al.* 2005; Chaudhary *et al.* 2016, 2017). Previous tests in our simulation,

for example, showed that increasing the right skewness of the range size frequency distribution results in a bimodal diversity gradient (Fig. S8), though this pattern disappears when species preferentially originate in the tropics (Jablonski 1993; Allen & Gillooly 2006; Jablonski *et al.* 2006). Whatever the true diversity pattern is, biases in the distribution of historical sampling effort imposes a serious limitation to accurately describe this pattern. Despite recent evidences of a bimodal diversity gradient (Brayard *et al.* 2005; Tittensor *et al.* 2010; Chaudhary *et al.* 2016, 2017; Woolley *et al.* 2016), it seems that a unimodal pattern would be hardly detected with the digital information available today, especially in the open ocean. Therefore, our study reinforces that until we move the sampling effort to tropical waters and improve public repositories, we may be hypothesizing ecological and evolutionary processes to explain sampling bias, instead of true natural phenomena. Meanwhile, under-sampling bias should be considered as a null hypothesis (Coddington *et al.* 2009). We recommend integration of the historical distribution of sampling effort with null models as a cautionary step in analysis of marine ecological patterns.

Acknowledgements

A.M. is supported by a Ph.D. scholarship provided by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. T.F.R. has been continuously supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, grant PQ309550/2015-7). This project is also supported by INCT in Ecology, Evolution and Biodiversity Conservation, funded by MCTIC/CNPq (grant 465610/2014-5) and FAPEG (grant 201810267000023).

References

- Allen, A.P. & Gillooly, J.F. (2006). Assessing latitudinal gradients in speciation rates and biodiversity at the global scale. *Ecol. Lett.*, 9, 947–954.
- Bates, A.E., Bird, T.J., Stuart-Smith, R.D., Wernberg, T., Sunday, J.M., Barrett, N.S., *et al.* (2015). Distinguishing geographical range shifts from artefacts of detectability and sampling effort. *Divers. Distrib.*, 21, 13–22.
- Boussarie, G., Bakker, J., Wangensteen, O.S., Mariani, S., Bonnin, L., Juhel, J.-B., *et al.* (2018). Environmental DNA illuminates the dark diversity of sharks. *Sci. Adv.*, 4, eaap9661.
- Brayard, A., Escarguel, G. & Bucher, H. (2005). Latitudinal gradient of taxonomic richness: combined outcome of temperature and geographic mid-domains effects? *J. Zool. Syst. Evol. Res.*, 43, 178–188.
- Carranza, A., Colwell, R.K. & Rangel, T.F.L.V.B. (2008). Distribution of megabenthic gastropods along environmental gradients: the mid-domain effect and beyond. *Mar. Ecol. Prog. Ser.*, 367, 193–202.
- Chao, A. (1984). Non-parametric estimation of the number of classes in a population. *Scand. J. Stat.*, 11, 265–270.
- Chao, A. & Jost, L. (2012). Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology*, 93, 2533–2547.
- Chaudhary, C., Saeedi, H. & Costello, M.J. (2016). Bimodality of latitudinal gradients in marine species richness. *Trends Ecol. Evol.*, 31, 670–676.
- Chaudhary, C., Saeedi, H. & Costello, M.J. (2017). Marine species richness is bimodal with latitude: a reply to Fernandez and Marques. *Trends Ecol. Evol.*, 32, 234–237.
- Chiarucci, A., Enright, N.J., Perry, G.L.W., Miller, B.P. & Lamont, B.B. (2003). Performance of nonparametric species richness estimators in a high diversity plant community. *Divers. Distrib.*, 9, 283–295.
- Clarke, A. & Lidgard, S. (2000). Spatial patterns of diversity in the sea: bryozoan species richness in the North Atlantic. *J. Anim. Ecol.*, 69, 799–814.
- Coddington, J.A., Agnarsson, I., Miller, J.A., Kuntner, M. & Hormiga, G. (2009). Undersampling bias: The null hypothesis for singleton species in tropical arthropod surveys. *J. Anim. Ecol.*, 78, 573–584.
- Coelho, M.T.P., Dambros, C., Rosauer, D.F., Pereira, E.B. & Rangel, T.F. (2018). Effects of neutrality and productivity on mammal richness and evolutionary history in

- Australia. *Ecography*, in press.
- Colwell, R.K. & Coddington, J.A. (1994). Estimating terrestrial biodiversity through extrapolation. *Philos. Trans. R. Soc. B Biol. Sci.*, 311, 101–118.
- Colwell, R.K. & Hurtt, G.C. (1994). Nonbiological gradients in species richness and a spurious Rapoport effect. *Am. Nat.*, 144, 570–595.
- Costello, M.J., Coll, M., Danovaro, R., Halpin, P., Ojaveer, H. & Miloslavich, P. (2010). A census of marine biodiversity knowledge, resources, and future challenges. *PLoS One*, 5, e12110.
- Fernandez, M.O. & Marques, A.C. (2017). Diversity of diversities: a response to Chaudhary, Saeedi, and Costello. *Trends Ecol. Evol.*, 32, 232–234.
- Fischer, A.G. (1960). Latitudinal variations in organic diversity. *Evolution*, 14, 64–81.
- Gaston, K.J. (1996). Species-range-size distributions: patterns, mechanisms and implications. *Trends Ecol. Evol.*, 11, 197–201.
- Hillebrand, H. (2004). On the generality of the latitudinal diversity gradient. *Am. Nat.*, 163, 192–211.
- Horn, M.H. & Allen, L.G. (1978). A distributional analysis of California coastal marine fishes. *J. Biogeogr.*, 5, 23–42.
- Hortal, J., de Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M. & Ladle, R.J. (2015). Seven shortfalls that beset large-scale knowledge of biodiversity. *Annu. Rev. Ecol. Evol. Syst.*, 46, 523–549.
- Jablonski, D. (1993). The tropics as a source of evolutionary novelty through geological time. *Nature*, 364, 142–144.
- Jablonski, D., Kaustuv, R. & Valentine, J.W. (2006). Out of the tropics: evolutionary diversity gradient. *Science*, 314, 102–106.
- Jetz, W. & Rahbek, C. (2001). Geometric constraints explain much of the species richness pattern in African birds. *Proc. Natl. Acad. Sci.*, 98, 5661–5666.
- Karenzi, N., Sink, K., Nel, R., Clark, A.E. & Altwegg, R. (2018). Imperfect detection distorts depth-related trends in marine macrofaunal species richness. *Ecography*, 41, 1698–1706.
- Kinlock, N.L., Prowant, L., Herstoff, E.M., Foley, C.M., Akin-Fajiye, M., Bender, N., *et al.* (2018). Explaining global variation in the latitudinal diversity gradient: meta-analysis confirms known patterns and uncovers new ones. *Glob. Ecol. Biogeogr.*, 27, 125–141.

- Lopez, L.C.S., De Aguiar Fracasso, M.P., Mesquita, D.O., Palma, A.R.T. & Riul, P. (2012). The relationship between percentage of singletons and sampling effort: A new approach to reduce the bias of richness estimates. *Ecol. Indic.*, 14, 164–169.
- MacNeil, M.A., Tyler, E.H.M., Fonnesebeck, C.J., Rushton, S.P., Polunin, N.V.C. & Conroy, M.J. (2008). Accounting for detectability in reef-fish biodiversity estimates. *Mar. Ecol. Prog. Ser.*, 367, 249–260.
- Menegotto, A. & Rangel, T.F. (2018). Mapping knowledge gaps in marine diversity reveals a latitudinal gradient of missing species richness. *Nat. Commun.*, 9, 4713.
- Meyer, C., Kreft, H., Guralnick, R. & Jetz, W. (2015). Global priorities for an effective information basis of biodiversity distributions. *Nat. Commun.*, 6, 8221.
- Mora, C., Tittensor, D.P. & Myers, R.A. (2008). The completeness of taxonomic inventories for describing the global diversity and distribution of marine fishes. *Proc. R. Soc. B Biol. Sci.*, 275, 149–155.
- O’Hara, T.D., Rowden, A.A. & Bax, N.J. (2011). A Southern Hemisphere bathyal fauna is distributed in latitudinal bands. *Curr. Biol.*, 21, 226–230.
- Peck, S.L. (2004). Simulation as experiment: a philosophical reassessment for biological modeling. *Trends Ecol. Evol.*, 19, 530–534.
- R Development Core Team (2018). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rangel, T.F., Edwards, N.R., Holden, P.B., Diniz-Filho, J.A.F., Gosling, W.D., Coelho, M.T.P., *et al.* (2018). Modeling the ecology and evolution of biodiversity: biogeographical cradles, museums, and graves. *Science*, 361, eaar5452.
- Rangel, T.F.L.V.B., Diniz-Filho, J.A.F. & Colwell, R.K. (2007). Species richness and evolutionary niche dynamics: a spatial pattern-oriented simulation experiment. *Am. Nat.*, 170, 602–616.
- Sanders, H.L. (1968). Marine benthic diversity: a comparative study. *Am. Nat.*, 102, 243–282.
- Stehli, F.G., McAlester, A.L. & Helsley, C.E. (1967). Taxonomic diversity of recent bivalves and some implications for geology. *Geol. Soc. Am. Bull.*, 78, 455–466.
- Thorson, G. (1957). Bottom communities (sublittoral or shallow Shelf). In: *Treatise on marine ecology and palaeoecology* (ed. Hedgpeth, J.W.). Geological Society of America, Baltimore, pp. 461–534.
- Tittensor, D., Mora, C., Jetz, W., Lotze, H., Ricard, D., Vanden Berghe, E., *et al.* (2010).

- Global patterns and predictors of marine biodiversity across taxa. *Nature*, 466, 1098–1101.
- Tittensor, D.P. & Worm, B. (2016). A neutral-metabolic theory of latitudinal biodiversity. *Glob. Ecol. Biogeogr.*, 25, 630–641.
- Webb, T.J., vanden Berghe, E. & O’Dor, R. (2010). Biodiversity’s big wet secret: the global distribution of marine biological records reveals chronic under-exploration of the deep pelagic ocean. *PLoS One*, 5, e10223.
- Whittaker, R.J., Araujo, M.B., Jepson, P., Ladle, R.J., Watson, J.E.M. & Willis, K.J. (2005). Conservation biogeography: assessment and prospect. *Divers. Distrib.*, 11, 3–23.
- Woolley, S.N.C., Tittensor, D.P., Dunstan, P.K., Guillera-Arroita, G., Lahoz-Monfort, J.J., Wintle, B.A., *et al.* (2016). Deep-sea diversity patterns are shaped by energy availability. *Nature*, 533, 393–396.

Considerações finais

A distribuição da biodiversidade marinha representa um dos componentes fundamentais para compreender os padrões globais de diversidade. No entanto, isso demanda um conhecimento minimamente eficiente sobre a distribuição das espécies. Por isso, instituições de países tropicais devem incentivar a pesquisa em regiões identificadas como menos investigadas, facilitando o deslocamento de especialistas entre bases internas e fortalecendo parcerias com instituições de países desenvolvidos para compartilhar estruturas de pesquisa que permitam ampliar a malha amostral em regiões onde a atividade de coleta pode ser tecnicamente desafiadora. Paralelamente, é extremamente importante incentivar a inclusão dos registros biológicos, contemporâneos e históricos, em redes internacionais dedicadas a disponibilização de dados científicos. Enquanto isso, estudos investigando gradientes de diversidade no ambiente marinho devem avaliar o fenômeno criticamente, explorando as possíveis causas da ausência de espécies nos trópicos e verificando a robustez dos métodos de estimativa diante do forte viés amostral. Por fim, é recomendável que ecólogos marinhos e terrestres ampliem suas redes profissionais de colaboração para além da interface terra-mar. O aumento no fluxo de hipóteses e dados descritivos entre sistemas traz consigo um potencial de grande generalidade, fortalecendo assim o crescimento conjunto da macroecologia.

APÊNDICES

Material suplementar

Capítulo 1

Appendix S1. Size and spatial distribution of islands and archipelagos

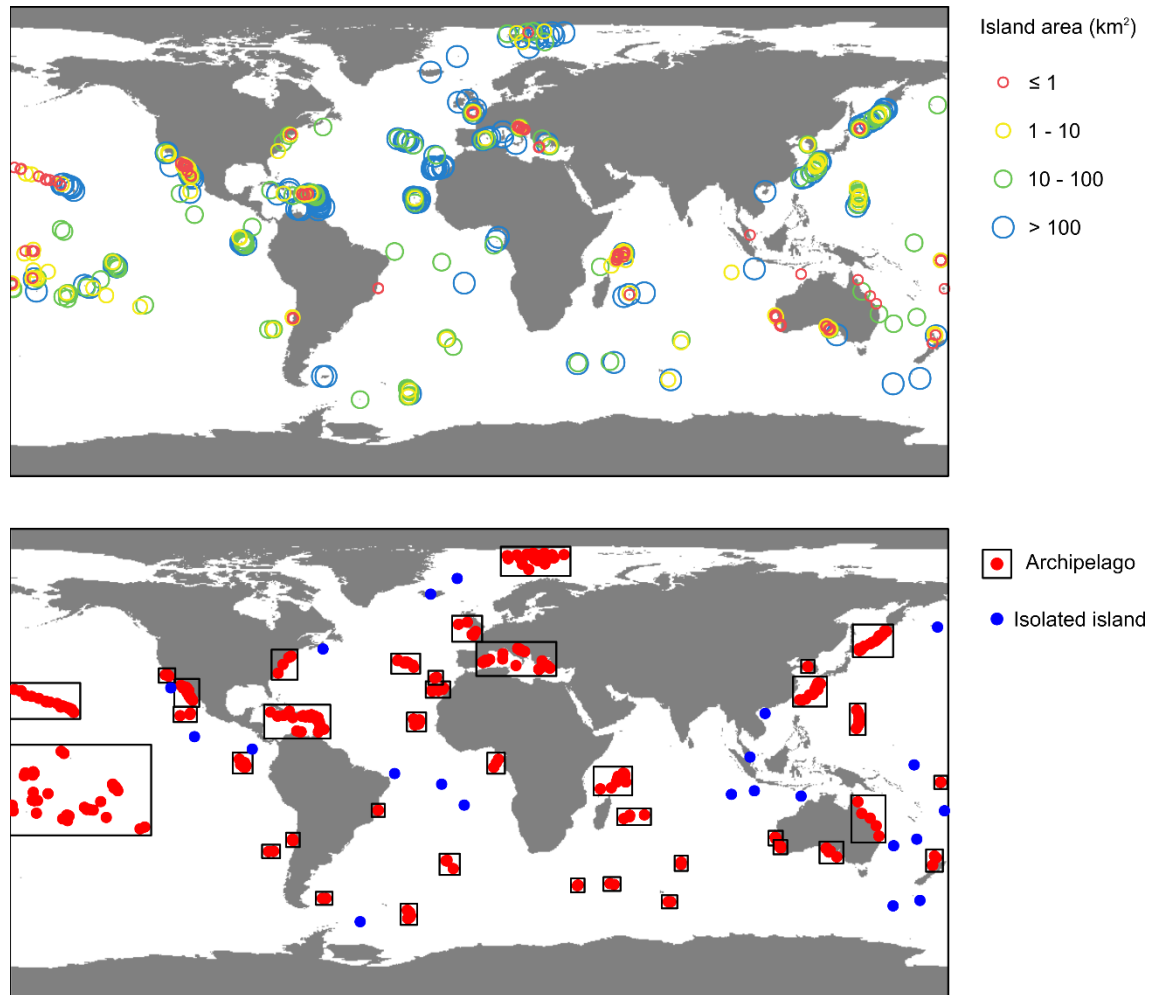


FIGURE S1. Spatial distribution of 790 islands used in this study. The size of the islands is shown for the centroid of the island through four categories (top). The distribution of archipelagos or set of close islands is shown with black squares (bottom). Red circles represent the islands within the archipelagos; blue circles represent isolated islands in our dataset.

Appendix S2. Comparing NPP models and raster resolution

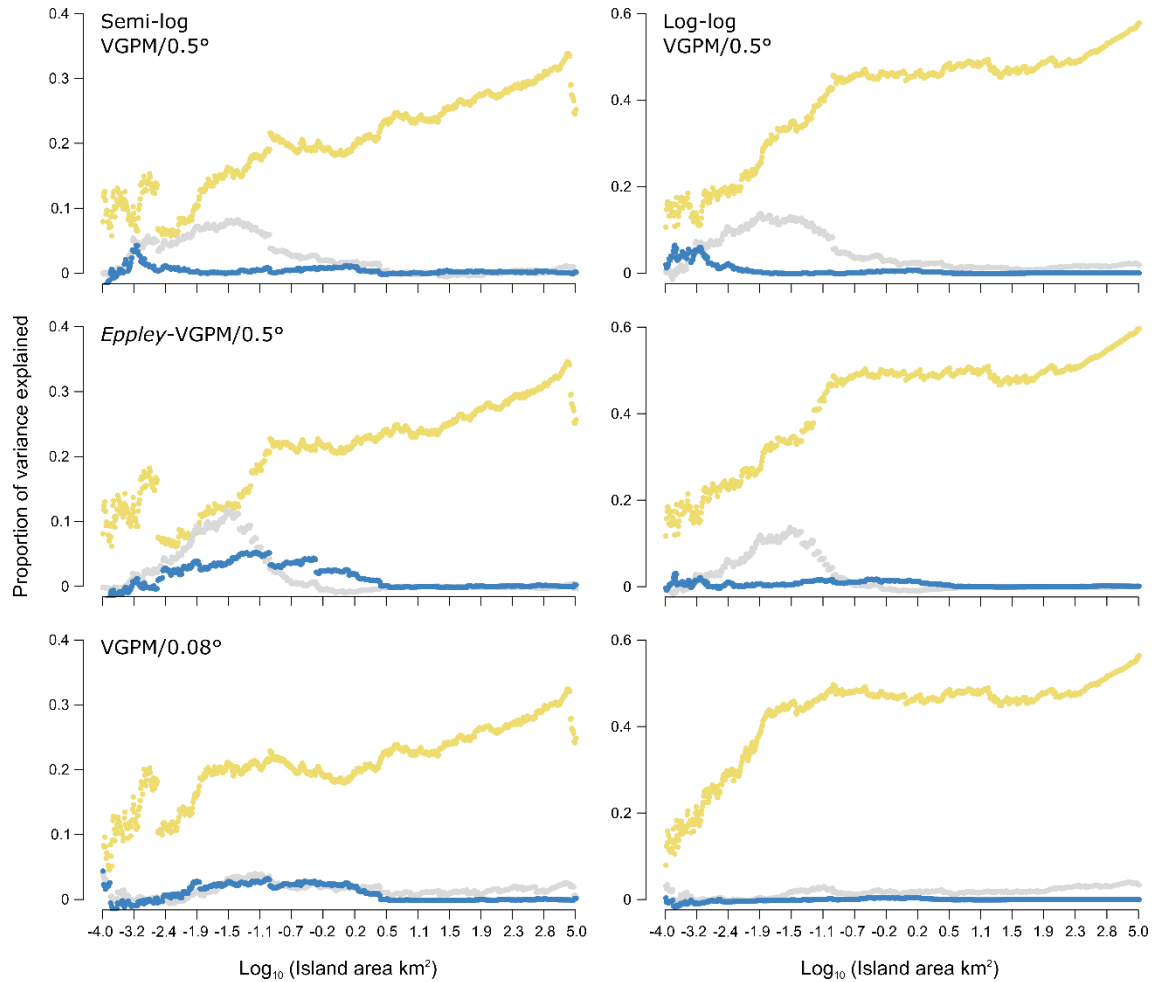


FIGURE S2. Variance in angiosperm species richness explained by island area (yellow dots) and marine NPP (blue dots) in a multiple linear regression. Plots show the exclusive contribution of each explanatory variable and the combined fraction (grey dots) as the size of the islands increases. Results are shown for three marine NPP models/resolutions and data transformations. Each tick in the x-axis represents the addition of 50 islands. Spatial resolution of NPP, or different models of productivity calculations (VGPM versus *Eppley*-VGPM), do not alter the main results of the variation partition analysis.

Appendix S3. Small islands threshold

We explored the most probable threshold in which insular species richness remains relatively independent of area, the upper limit of the small-island effect (SIE), through two breakpoint regression models: a continuous piecewise regression and a shallow-step zero slope piecewise regression. The continuous piecewise regression was calculated using the function ‘segmented’ implemented in the homonymous R package (Muggeo 2017). The function ‘segmented’ uses maximum likelihood to determine the most probable breakpoint in the relationship analyzed. The shallow-step zero slope piecewise regression was calculated using the equation proposed by Lomolino & Weiser (2001):

$$Y = c + z [(\log_{10}(A) > T) \times (\log_{10}(A) - T)]$$

where Y is species richness (S) or $\log_{10}(S)$ for semi-log and log-log transformations, respectively, A is island area in km^2 and T is the upper limit of the SIE. We built an algorithm in the R statistical programming language to run the equation considering each island size, from the smallest to the biggest, as a possible threshold. The algorithm returns the island size where the breakpoint model best fits the data (highest R^2), which we assume to be the ideal upper limit of the SIE.

The application of a semi-log transformation resulted in a pattern of the species-area relationship similar to the expected according to the concept of the SIE (Fig. S3a). However, the estimated breakpoint position differed largely between the two breakpoint models (continuous piecewise regression: 120 km^2 ; shallow-step piecewise regression: 35 km^2). Both models performed better than a simple linear regression (R^2_{adj} continuous: 0.46; R^2_{adj} shallow-step: 0.43; R^2 linear: 0.25). Using log-log transformation, there was no clear difference among the models (Fig. S3b). Continuous piecewise regression returned an

estimated breakpoint at 0.014 km², whereas for the shallow-step regression the model without breakpoint fitted better. The fraction of variance in angiosperm species richness explained by island area did not varied among the models (R^2_{adj} continuous: 0.6; R^2_{adj} shallow-step: 0.6; R^2 linear: 0.6).

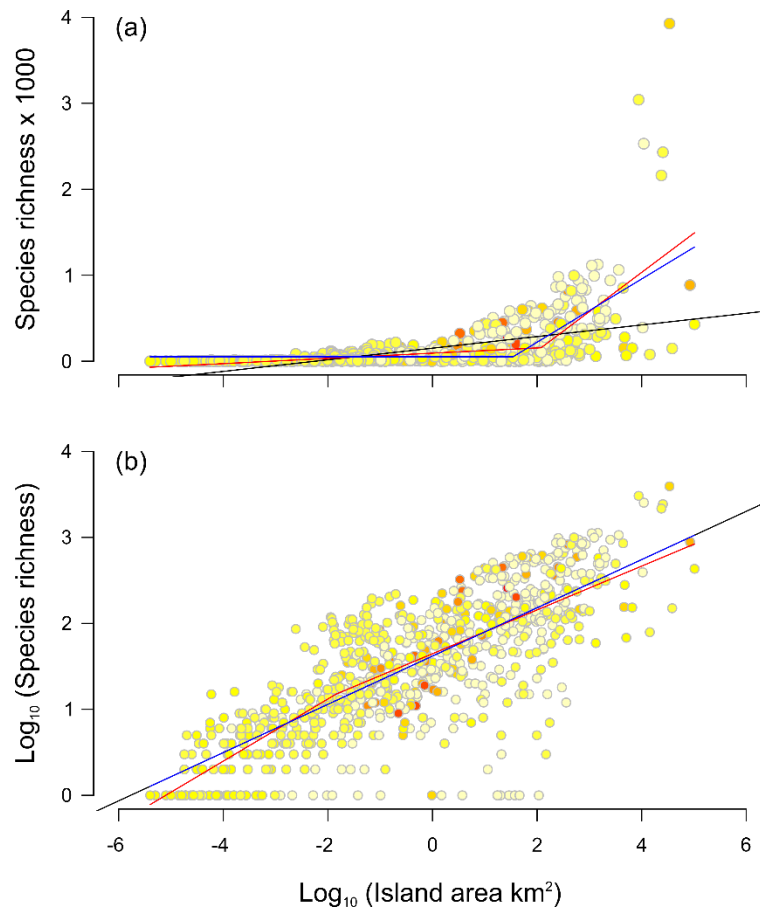


FIGURE S3. Species-area relationship for angiosperms of 790 islands distributed worldwide. Plots show results for two different data transformations: semi-log (a) and log-log (b). Lines represent the relationship between species richness and area considering a linear model (black line), a continuous piecewise regression (red lines) and a shallow-step piecewise regression (blue line). Colored circles show marine productivity around the island, with hot colors indicating higher marine productivity

Appendix S4. Collinearity among variables

Before applying the multiple regression analysis, we evaluated the correlation among log area, mean annual marine net primary production (NPP) and their interaction term. The results showed that the interaction term was highly correlated with log area ($r = 0.89$), indicating that collinearity may influence the analyses. Marine NPP, on the other hand, was weakly correlated with both log area ($r = -0.21$) and the interaction term ($r = -0.15$). The results of the variance inflation factor confirmed the collinearity between log area and the term of interaction (Table S1).

TABLE S1. Variance inflation factors (VIF) for log area (Area), marine net primary production (marine NPP) and the term of interaction in multiple regression for prediction of log angiosperm species richness in islands.

	Without interaction term	With interaction term
Area	1.04828	4.92292
Marine NPP	1.04828	1.05547
Area*Marine NPP	-	4.81018

Appendix S5. Influence of seabirds' presence

Marine subsidy for angiosperms of small islands is expected to be supplied by seabirds in the form of guano. For this reason, we might be adding noise in our estimates by including islands without the main subsidy vector in the analyses. To evaluate this potential influence, we repeated all analyses including only those islands where seabirds are present. We considered all species of the following clades to be seabirds: Sphenisciformes, Procellariiformes, Pelecaniformes, Phaethontiformes, Suliformes (without Anhingidae) and the Charadriiformes families Stercorariidae, Alcidae and Laridae (Schreiber & Burger 2002). We used two approaches to verify the presence of seabirds on each island. First, we retrieved range maps from BirdLife International and NatureServe (<http://datazone.birdlife.org/species/requestdis>; Version 7.0) of 492 species of seabirds (extant, native, resident or occurring in either breeding or non-breeding season) to identify which islands were within the occurrence range of any seabird species (the specific procedure is explained below). Because all islands in our dataset were within the range of at least one seabird species, we applied a second, more precise, approach by verifying the records of seabird observation from Global Biodiversity Information Facility (GBIF; <http://www.gbif.org>). We retrieved more than 60 million records (60,795,237) of seabirds from GBIF database at 30/01/2018. We filtered the data by eliminating all records of fossils, records without coordinates information or duplicated. After the quality control procedure, a total of 8,891,512 valid records for seabirds' presence across the world was used in our analyses. We used the 'over' function from the 'raster' R package (Hijmans 2018) to overlap the seabirds' presence records with the islands polygons to determine the islands where seabirds were present. This revealed that a total of 377 islands in our dataset, ranging from 0.24 km² to 102,000 km², have a presence of seabirds confirmed. As can be noted below

(Fig. S4-S7), the results of our study did not change when only those 377 islands were included in the analyses.

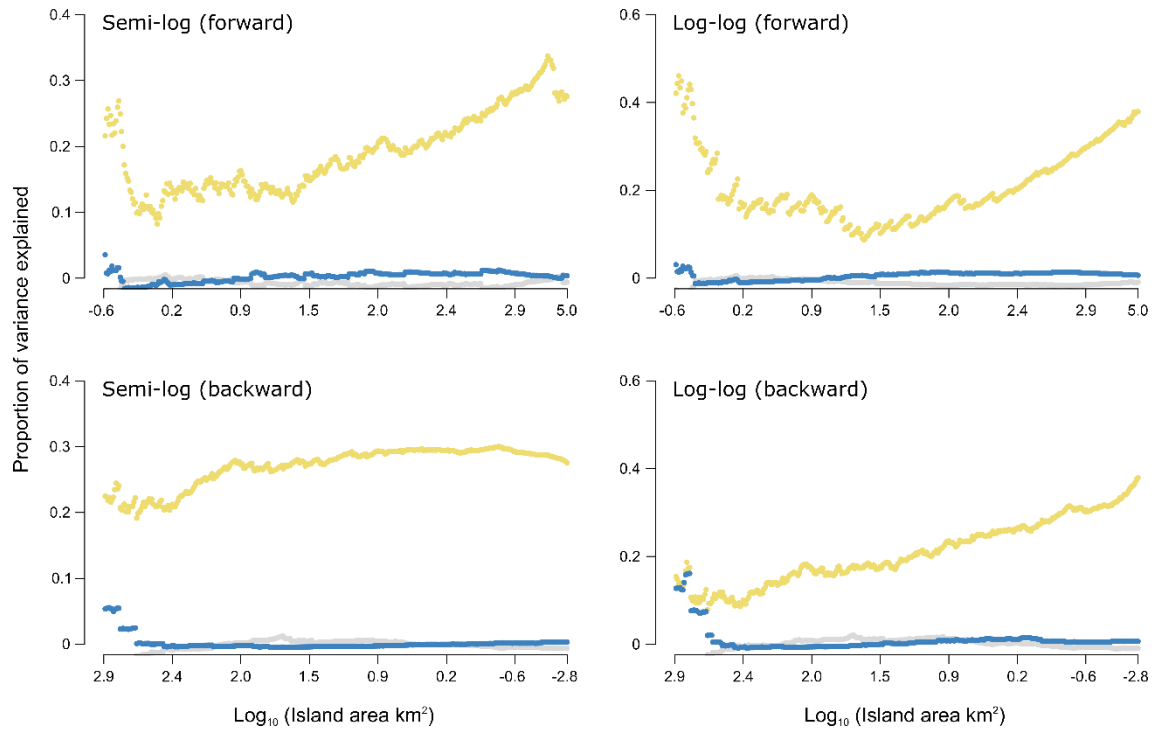


FIGURE S4. Variance in angiosperm species richness explained by island area (yellow dots) and marine NPP (blue dots) in a multiple linear regression considering only those islands with seabirds present. Plots show the exclusive contribution of each explanatory variable and the combined fraction (grey dots) as the size of the islands increases (top) or decreases (bottom). Each tick in the x-axis represents the addition of 50 islands.

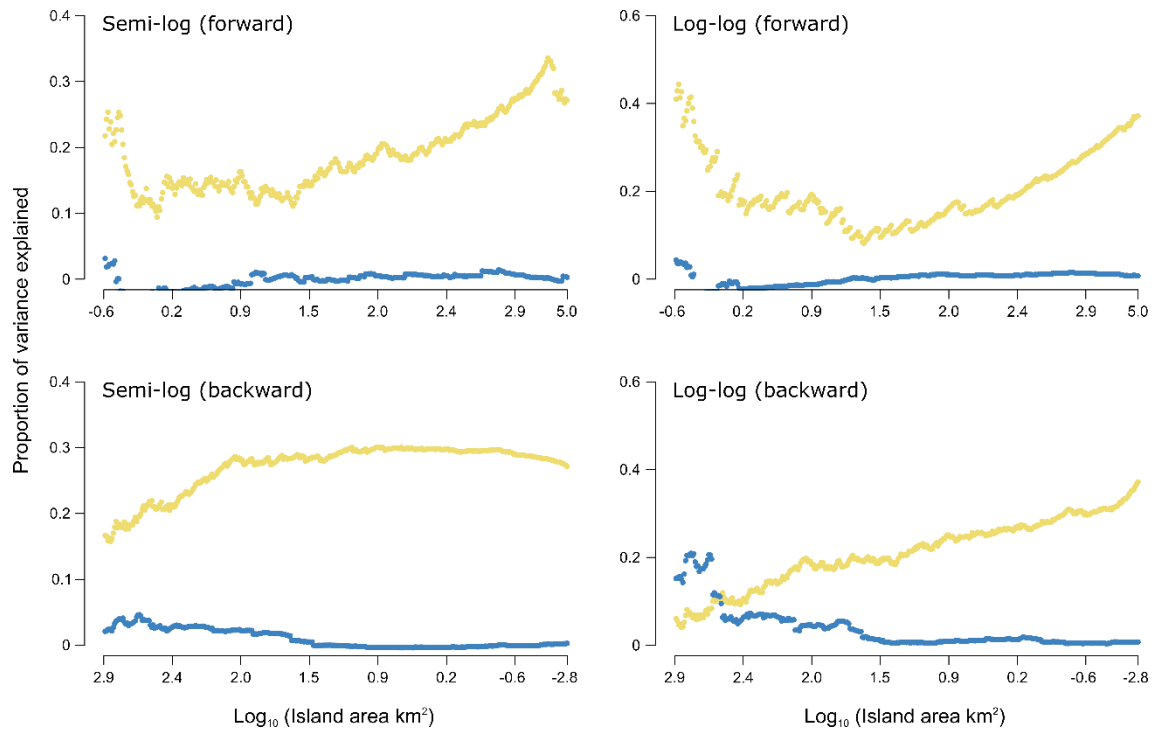


FIGURE S5. Variance in angiosperm species richness explained by island area (yellow dots) and marine NPP (blue dots) considering only those islands with seabirds present. Plots show the linear contribution of area and the quadratic effect of marine NPP on the residuals of the species-area relationship for insular angiosperms as the size of the islands increases (top) or decreases (bottom). Each tick in the x-axis represents the addition of 50 islands.

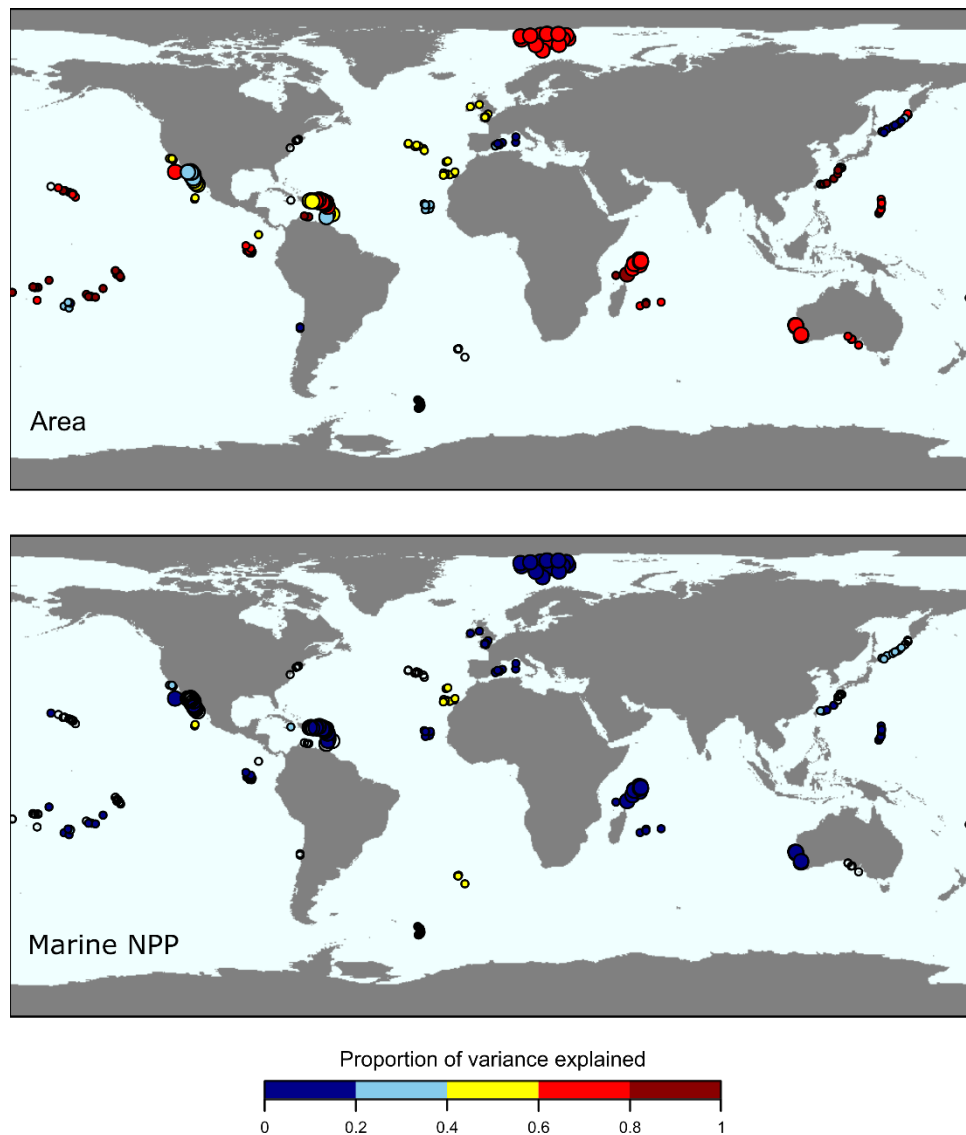


FIGURE S6. Non-stationarity in the effects of island area and marine NPP on angiosperms species richness considering only those islands with seabirds present. The maps show the spatial distribution of the fraction explained exclusively by area and marine NPP for each set of islands. Open circles represent non-estimated or abnormal (negative) R-squared values. To evaluate the influence of low sampling size, we plotted with small circles results of regressions calculated with less the 15 islands. Variation partitioning calculated with semi-log transformed data.

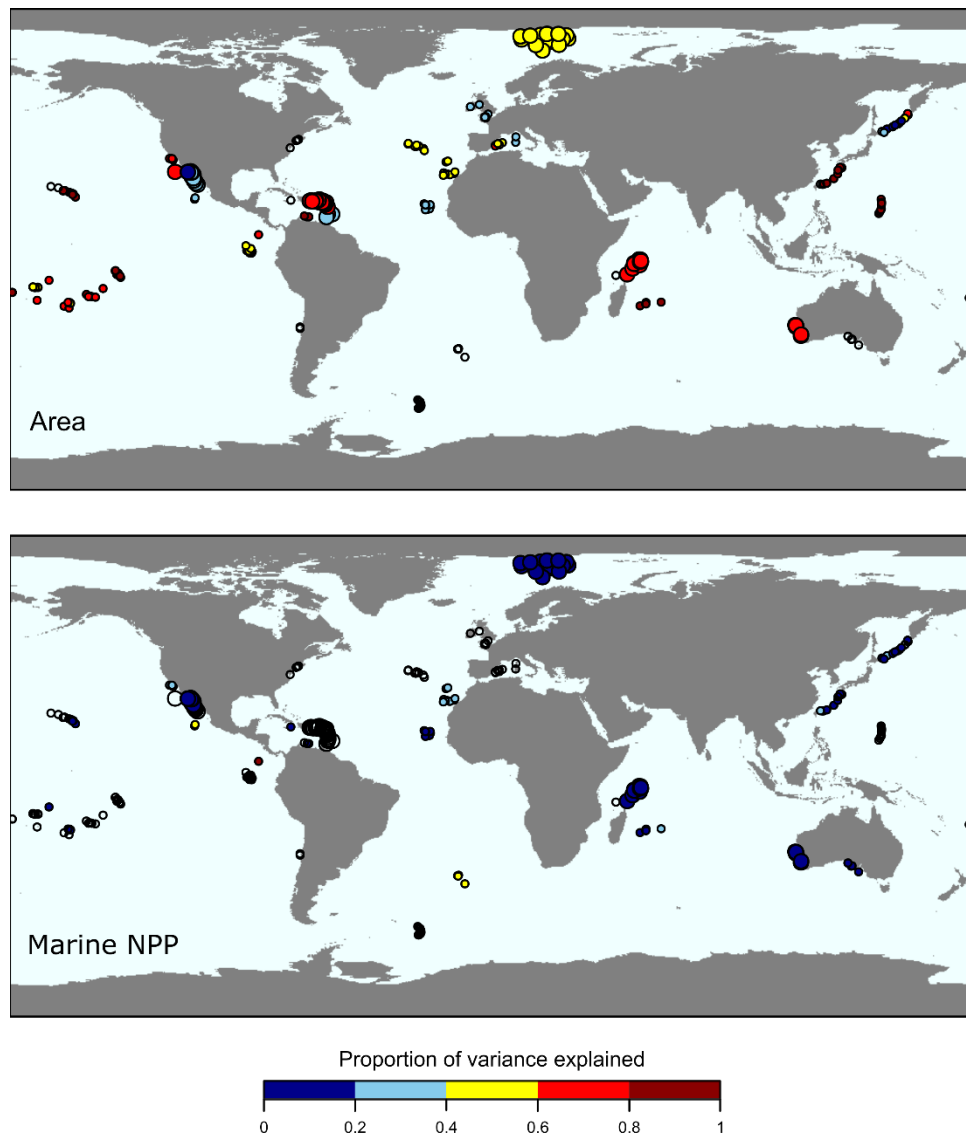


FIGURE S7. Non-stationarity in the effects of island area and marine NPP on angiosperms species richness considering only those islands with seabirds present. The maps show the spatial distribution of the fraction explained exclusively by area and marine NPP for each set of islands. Open circles represent non-estimated or abnormal (negative) R-squared values. To evaluate the influence of low sampling size, we plotted with small circles results of regressions calculated with less the 15 islands. Variation partitioning calculated with log-log transformed data.

Appendix S6. GWR for semi-log model

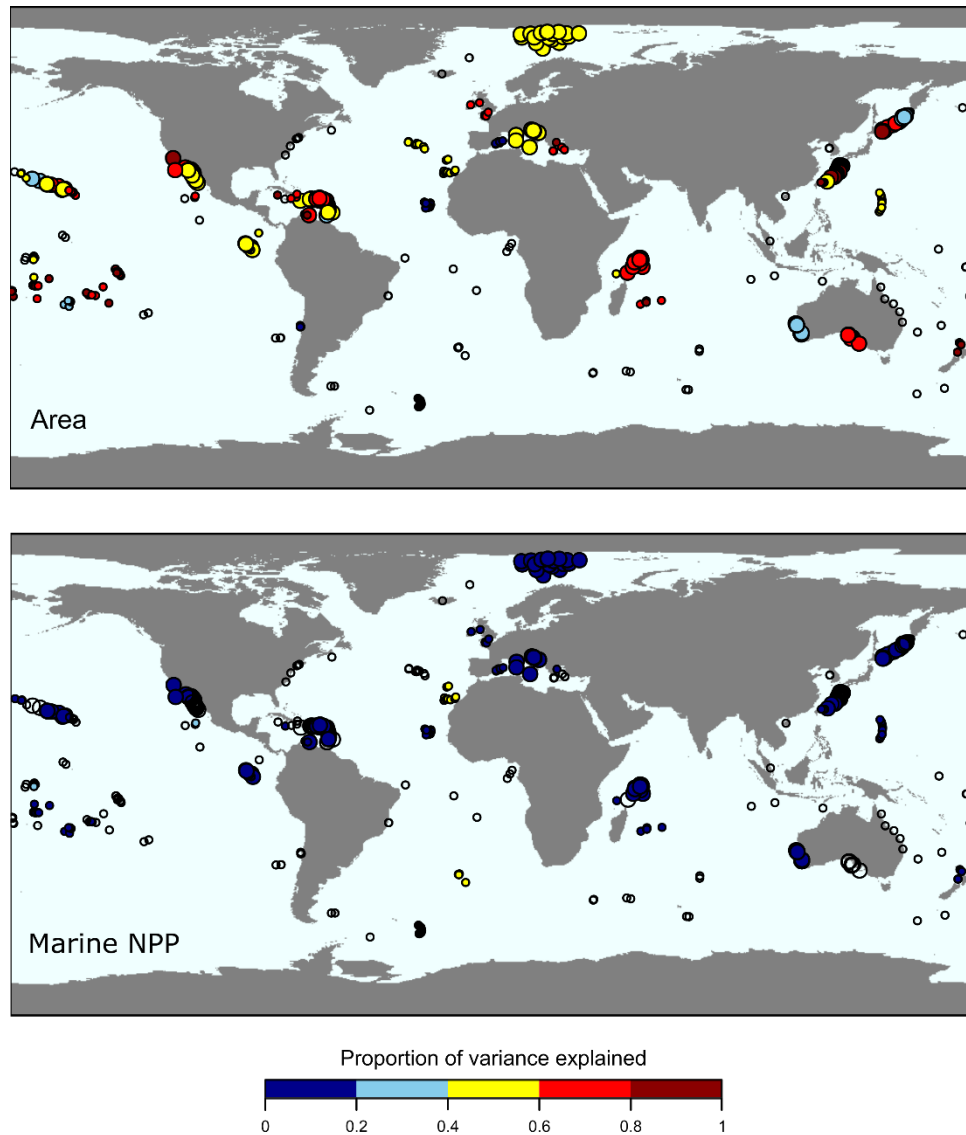


FIGURE S8. Non-stationarity in the effects of island area and marine NPP on angiosperms species richness. The maps show the spatial distribution of the fraction explained exclusively by area and marine NPP for each set of islands. Open circles represent non-estimated or abnormal (negative) R-squared values. To evaluate the influence of low sampling size, we plotted with small circles results of regressions calculated with less the 15 islands. Variation partitioning calculated with semi-log transformed data.

Appendix S7. Generalized linear mixed model results

Results of the generalized linear mixed model considering island area, distance to the coast, connection to the mainland during the last glacial maximum, temperature, precipitation, and marine NPP as the explanatory fixed variables and archipelago as the random variable. The results are shown for the entire dataset (Table S2) and considering only small islands, islands with seabirds, islands with at least one angiosperm species, and islands with seabirds and at least one angiosperm species (Table S3).

TABLE S2. Effects of island area (Area), marine productivity (NPP), distance to the coast (Dist), last glacial maximum mainland connection (GMMC), temperature (Temp) and precipitation (Prec) on insular angiosperm species richness shown by the generalized linear mixed model with archipelago as random variable.

Parameter	Estimate	SE	df	<i>t</i> -value	<i>P</i> value
Intercept	0.886	0.116	724	7.632	< 0.001
Area	0.340	0.009	724	38.438	< 0.001
NPP	-8.4E-06	4.2E-05	724	-0.201	0.8411
Dist	-1.8E-04	3.0E-05	724	-5.975	< 0.001
GMMC	0.256	0.051	724	4.993	< 0.001
Temp	0.024	0.005	724	5.156	< 0.001
Prec	2.4E-04	5.3E-05	724	4.603	< 0.001

Note. Species richness and area were log-transformed. See Fig. S1 for archipelagos distribution. There was no spatial autocorrelation in the residuals (Moran' *I* = -0.0002, *P* = 0.23).

TABLE S3. Effects of island area, marine productivity, distance to the coast, last glacial maximum mainland connection, temperature and precipitation on insular angiosperm species richness shown by the generalized linear mixed model with archipelago as random variable, considering alternative scenarios. See Table S2 for caption.

Parameter	Estimate	SE	df	<i>t</i> -value	<i>P</i> value
<i>Small islands (< 1 km²)</i>					
Intercept	1.173	0.260	392	4.509	< 0.001
Area	0.315	0.020	392	16.000	< 0.001
NPP	-3.9E-05	7.4E-05	392	-0.528	0.598
Dist	-1.5E-04	7.1E-05	392	-2.167	0.031
GMMC	0.335	0.103	392	3.263	0.001
Temp	0.001	0.010	392	0.114	0.909
Prec	3.8E-04	9.3E-05	392	4.059	< 0.001
<i>Islands with seabirds</i>					
Intercept	0.892	0.122	314	7.319	< 0.001
Area	0.325	0.012	314	28.104	< 0.001
NPP	-2.1E-05	4.5E-05	314	-0.469	0.640
Dist	-1.8E-04	3.1E-05	314	-5.764	< 0.001
GMMC	0.244	0.064	314	3.786	< 0.001
Temp	0.027	0.005	314	5.677	< 0.001
Prec	2.3E-04	5.9E-05	314	3.898	< 0.001
<i>Excluding empty islands</i>					
Intercept	0.988	0.111	674	8.882	< 0.001
Area	0.325	0.009	674	37.714	< 0.001
NPP	7.0E-07	4.0E-05	674	0.017	0.986
Dist	-1.8E-04	2.8E-05	674	-6.338	< 0.001
GMMC	0.210	0.048	674	4.326	< 0.001
Temp	0.021	0.004	674	4.750	< 0.001
Prec	2.3E-04	5.0E-05	674	4.653	< 0.001
<i>Islands with seabirds excluding empty islands</i>					
Intercept	0.943	0.116	304	8.120	< 0.001
Area	0.325	0.011	304	29.006	< 0.001
NPP	-6.5E-06	4.3E-05	304	-0.150	0.881
Dist	-1.8E-04	3.0E-05	304	-6.164	< 0.001
GMMC	0.168	0.064	304	2.629	0.009
Temp	0.026	0.005	304	5.692	< 0.001
Prec	2.2E-04	5.6E-05	304	4.028	< 0.001

Note. Species richness and area were log-transformed. See Fig. S1 for archipelagos distribution. There was no spatial autocorrelation in the residuals ($-0.0007 < \text{Moran}'I < -0.0001$, $P > 0.24$).

Appendix S8. Influence of empty islands

Results of our analyses considering only those islands with at least one angiosperm species, that is, excluding the empty islands (Fig. S9 – S12).

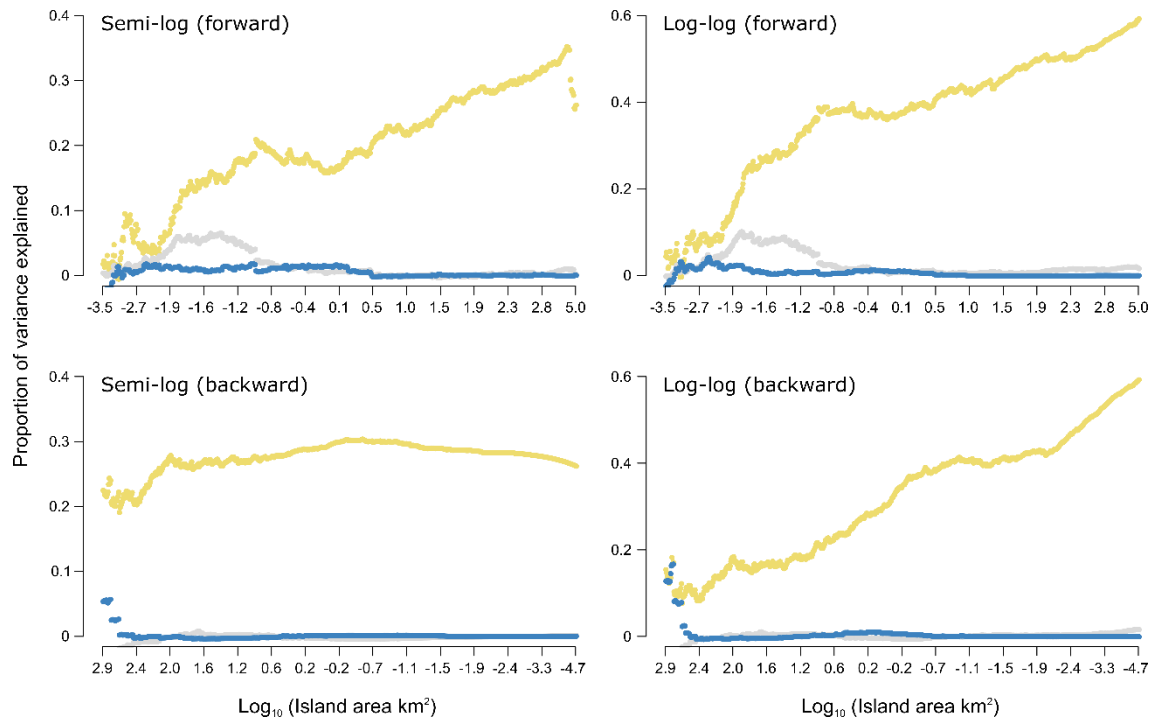


FIGURE S9. Variance in angiosperm species richness explained by island area (yellow dots) and marine NPP (blue dots) in a multiple linear regression considering only those islands with at least one angiosperm species. Plots show the exclusive contribution of each explanatory variable and the combined fraction (grey dots) as the size of the islands increases (top) or decreases (bottom). Each tick in the x-axis represents the addition of 50 islands.

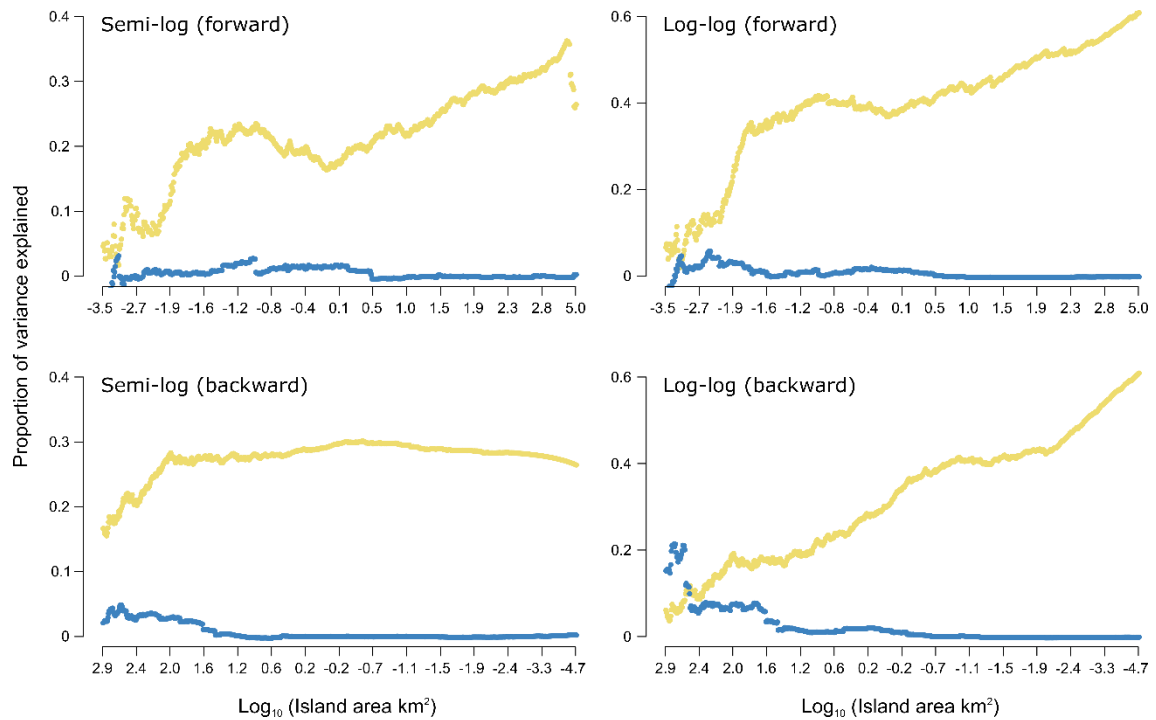


FIGURE S10. Variance in angiosperm species richness explained by island area (yellow dots) and marine NPP (blue dots) considering only those islands with at least one angiosperm species. Plots show the linear contribution of area and the quadratic effect of marine NPP on the residuals of the species-area relationship for insular angiosperms as the size of the islands increases (top) or decreases (bottom). Each tick in the x-axis represents the addition of 50 islands.

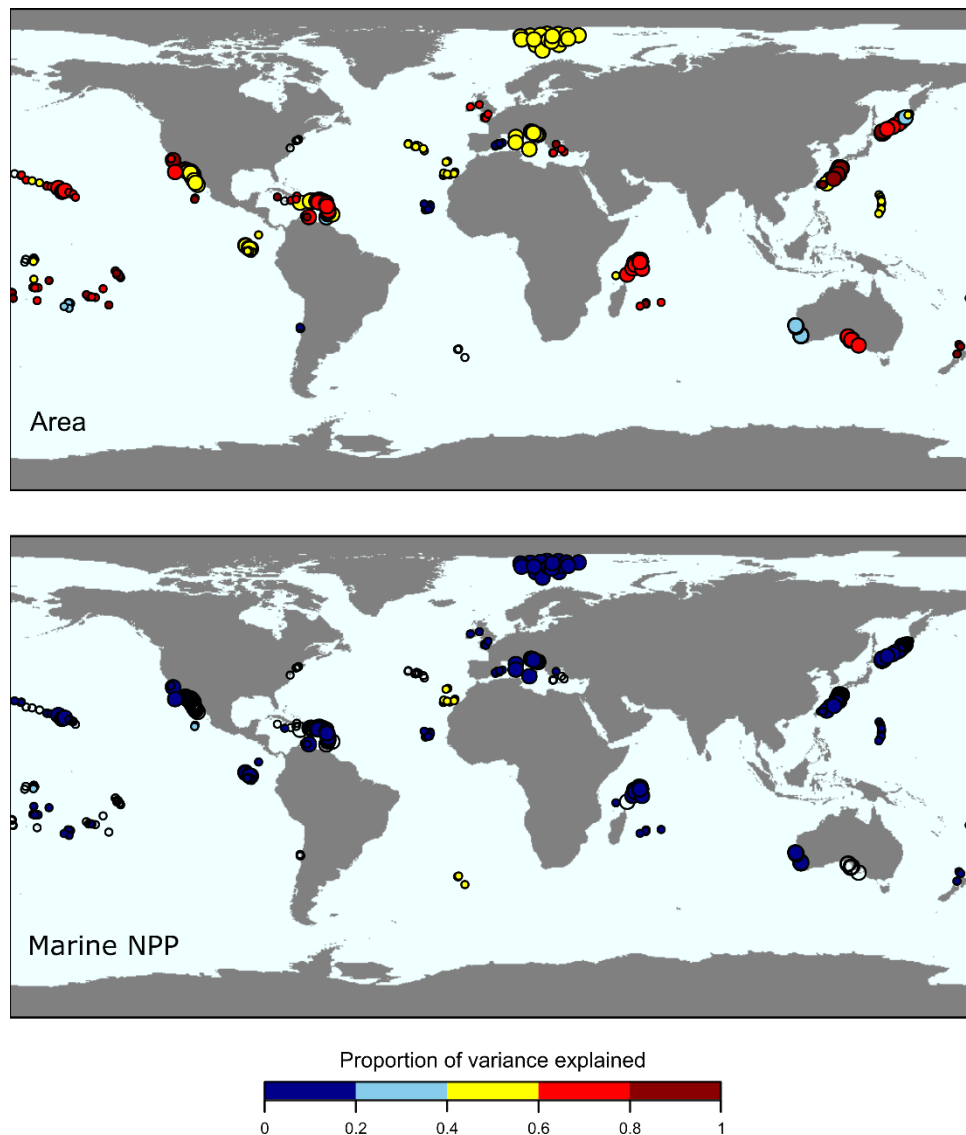


FIGURE S11. Non-stationarity in the effects of island area and marine NPP on angiosperms species richness considering only those islands with at least one species. The maps show the spatial distribution of the fraction explained exclusively by area and marine NPP for each set of islands. Open circles represent non-estimated or abnormal (negative) R-squared values. To evaluate the influence of low sampling size, we plotted with small circles results of regressions calculated with less the 15 islands. Variation partitioning calculated with semi-log transformed data.

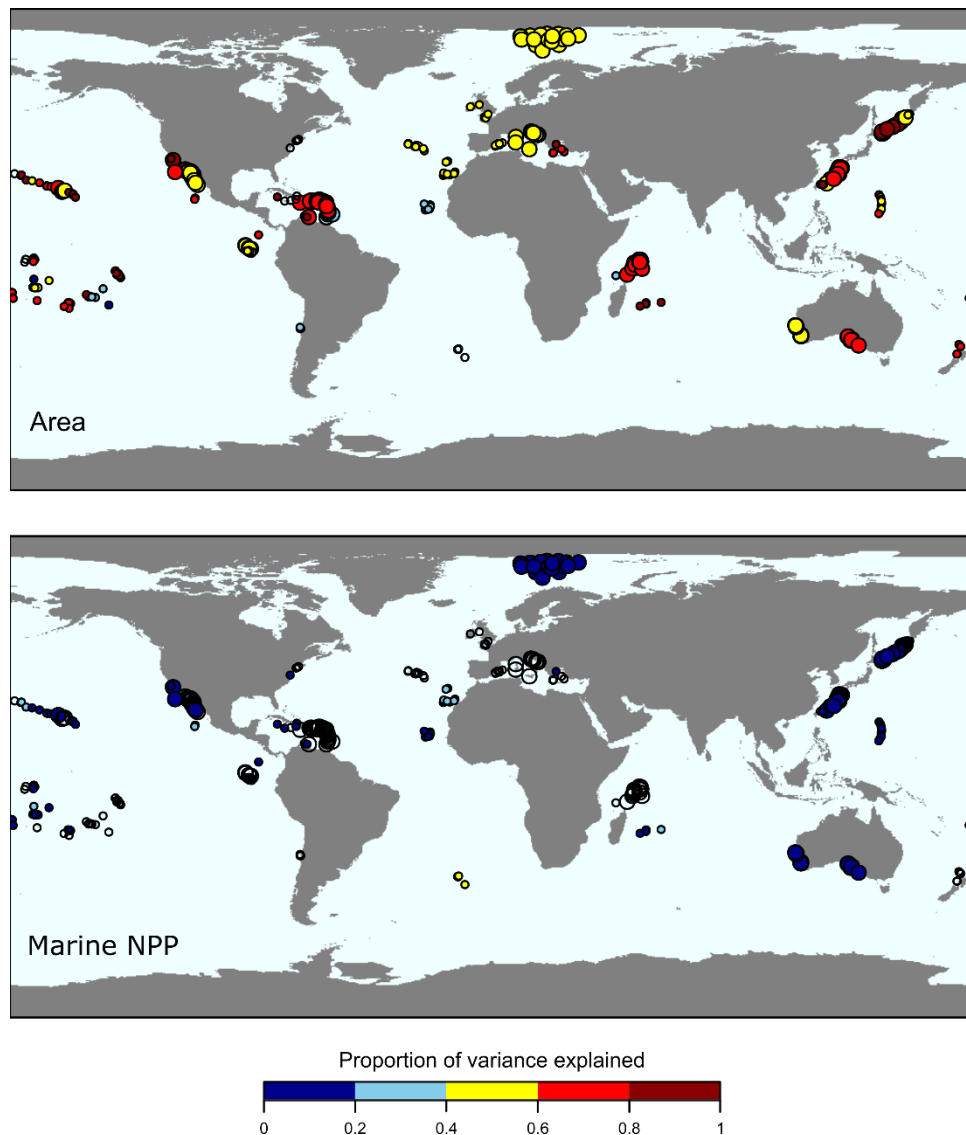


FIGURE S12. Non-stationarity in the effects of island area and marine NPP on angiosperms species richness considering only those islands with at least one species. The maps show the spatial distribution of the fraction explained exclusively by area and marine NPP for each set of islands. Open circles represent non-estimated or abnormal (negative) R-squared values. To evaluate the influence of low sampling size, we plotted with small circles results of regressions calculated with less the 15 islands. Variation partitioning calculated with log-log transformed data.

Appendix S9. Influence of empty islands and seabirds' presence

Results of our analyses considering only those islands with at least one angiosperm species and confirmed presence of seabirds (Fig. S13 – S16).

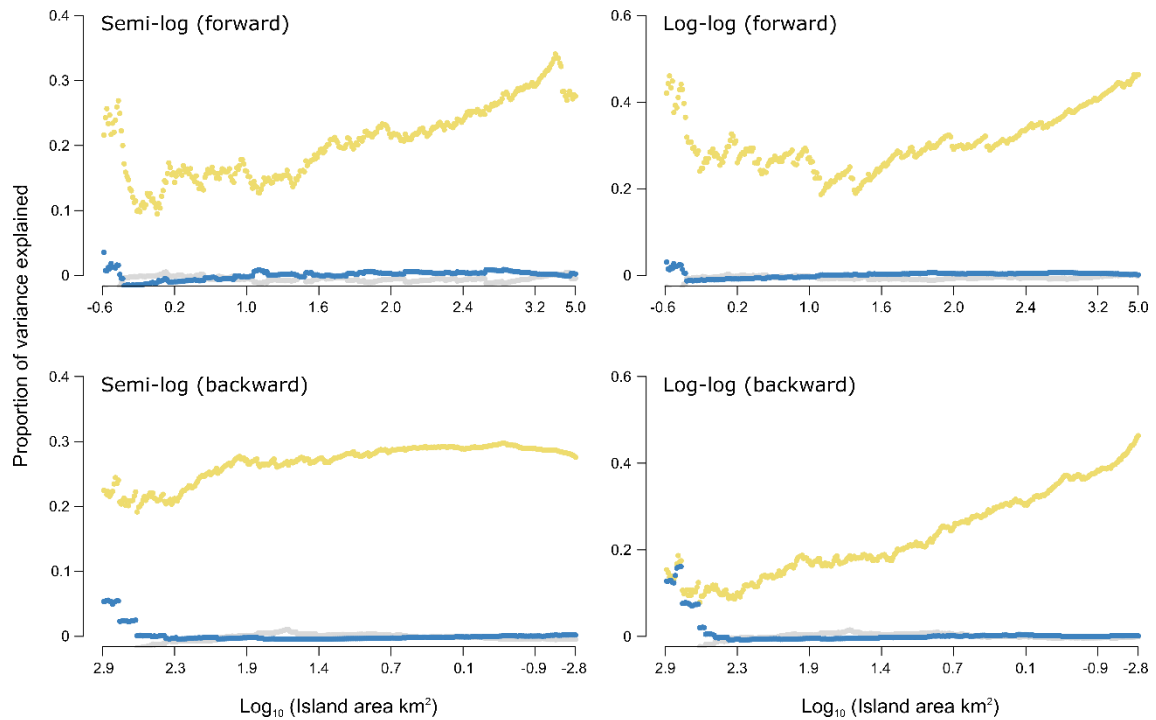


FIGURE S13. Variance in angiosperm species richness explained by island area (yellow dots) and marine NPP (blue dots) in a multiple linear regression considering only those islands with at least one angiosperm species and seabirds present. Plots show the exclusive contribution of each explanatory variable and the combined fraction (grey dots) as the size of the islands increases (top) or decreases (bottom). Each tick in the x-axis represents the addition of 50 islands.

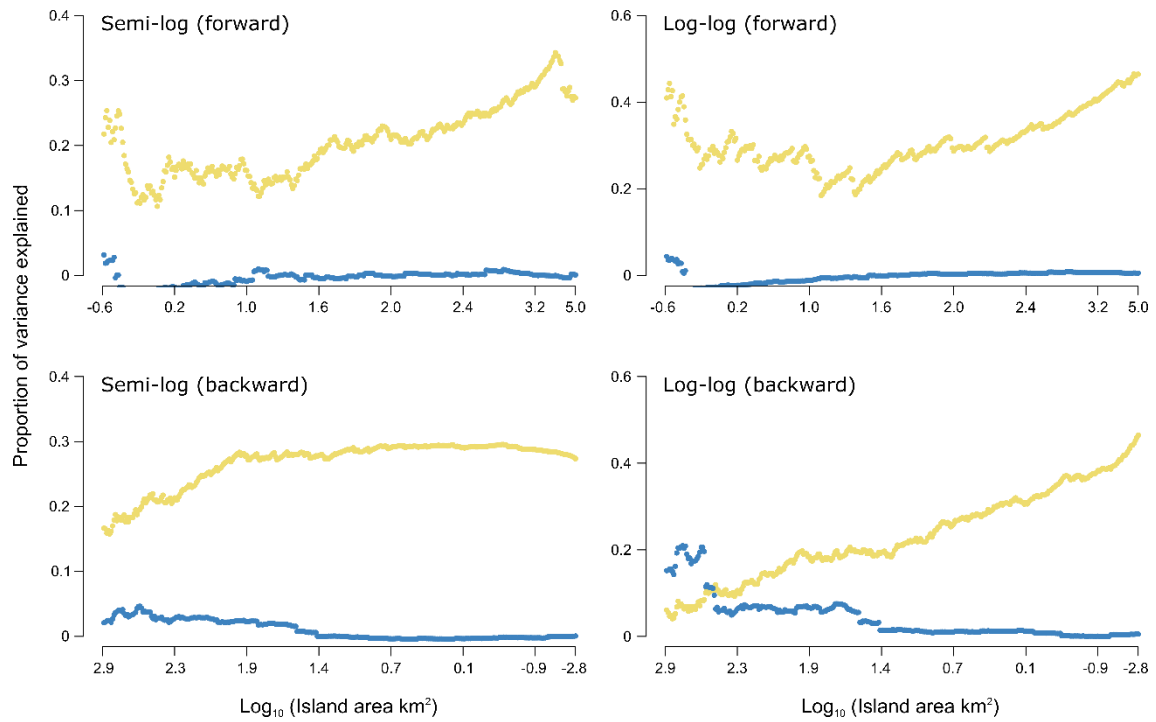


FIGURE S14. Variance in angiosperm species richness explained by island area (yellow dots) and marine NPP (blue dots) considering only those islands with at least one angiosperm species and seabirds present. Plots show the linear contribution of area and the quadratic effect of marine NPP on the residuals of the species-area relationship for insular angiosperms as the size of the islands increases (top) or decreases (bottom). Each tick in the x-axis represents the addition of 50 islands.

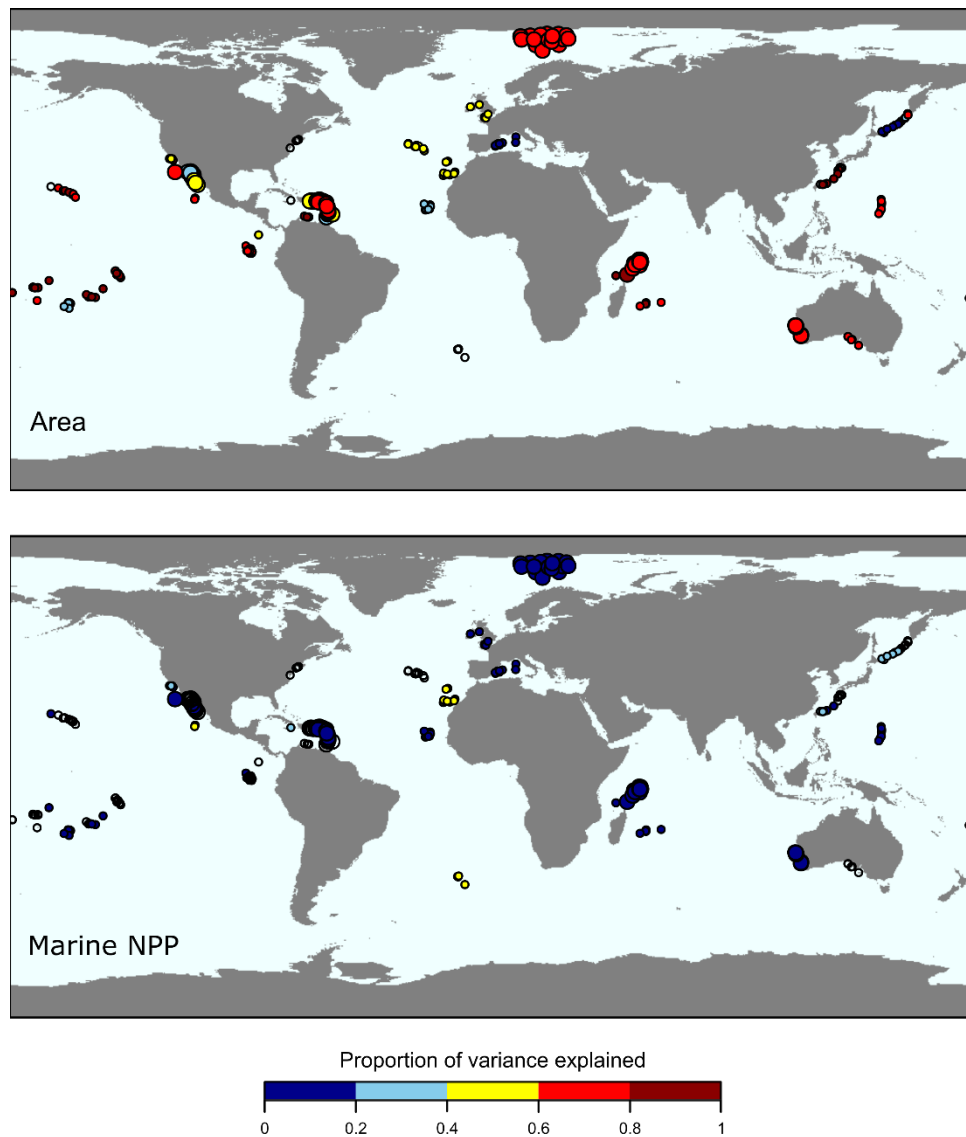


FIGURE S15. Non-stationarity in the effects of island area and marine NPP on angiosperms species richness considering only those islands with at least one angiosperm species and seabirds present. The maps show the spatial distribution of the fraction explained exclusively by area and marine NPP for each set of islands. Open circles represent non-estimated or abnormal (negative) R-squared values. To evaluate the influence of low sampling size, we plotted with small circles results of regressions calculated with less the 15 islands. Variation partitioning calculated with semi-log transformed data.

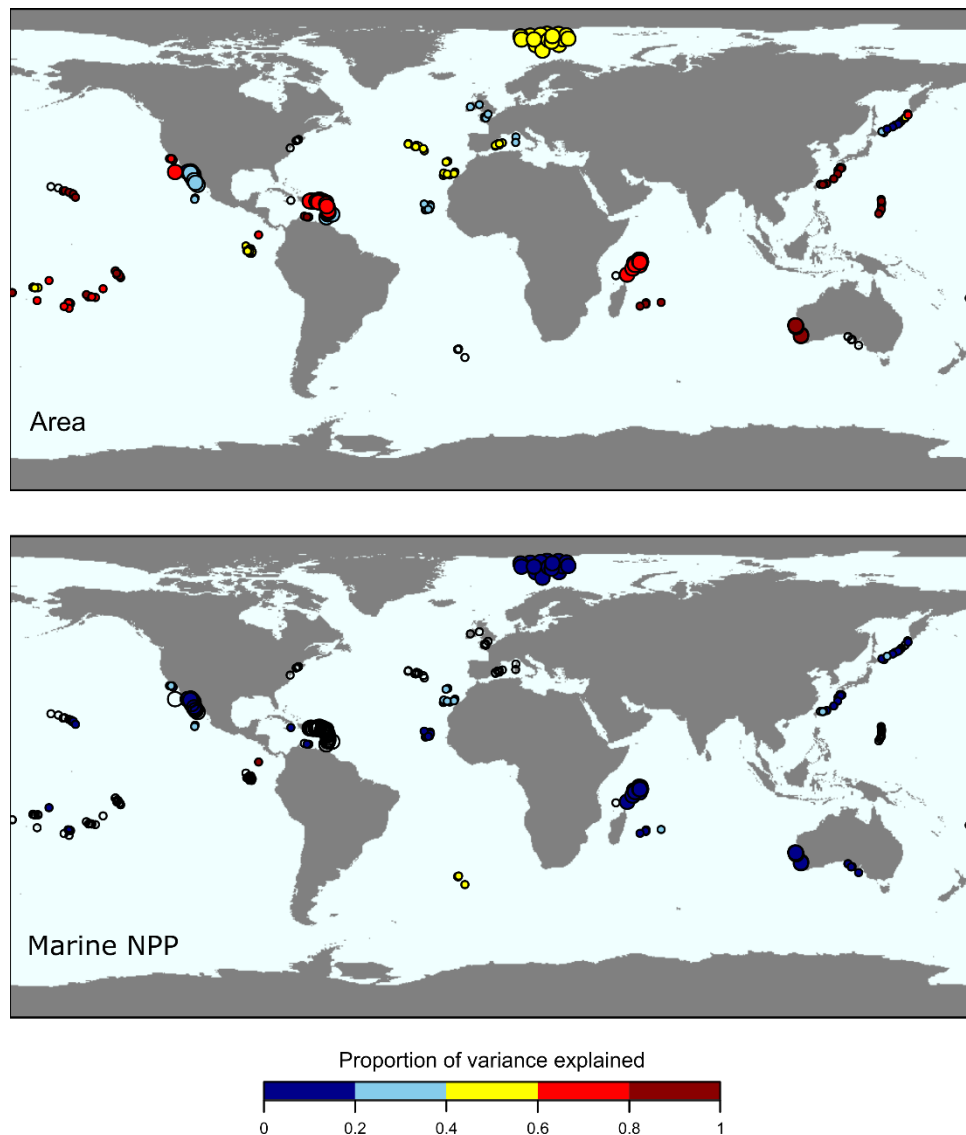


FIGURE S16. Non-stationarity in the effects of island area and marine NPP on angiosperms species richness considering only those islands with at least one angiosperm species and seabirds present. The maps show the spatial distribution of the fraction explained exclusively by area and marine NPP for each set of islands. Open circles represent non-estimated or abnormal (negative) R-squared values. To evaluate the influence of low sampling size, we plotted with small circles results of regressions calculated with less the 15 islands. Variation partitioning calculated with log-log transformed data.

Supplementary References

- Hijmans, R.J. (2018). raster: geographic data analysis and modeling. R package version 2.8-4. Available at: <https://CRAN.R-project.org/package=raster>
- Lomolino, M. V. & Weiser, M.D. (2001). Towards a more general species-area relationship: diversity on all islands, great and small. *J. Biogeogr.*, 28, 431–445.
- Muggeo, V.M.R. (2017). Segmented: regression models with break-points/change-points estimation. R package version 0.5-3.0. Available at: <https://CRAN.R-project.org/package=segmented>
- Schreiber, E.A. & Burger, J. (2002). *Biology of marine birds*. CRC Press, Boca Raton.

Material suplementar

Capítulo 3

Supplementary Note 1 | Latitudinal range of marine species

To evaluate the range contiguity assumption for marine species at the scale of our analysis (5° latitudinal bands), we quantified spatial gaps in the latitudinal distribution of relatively well-known species using range maps provided by the International Union for the Conservation of Nature (IUCN). The IUCN range maps are based not only on point-occurrence records, but also on expert knowledge of the biology of species and habitat preference. We measured latitudinal gaps in geographical distribution of all Scombridae, Elasmobranchii and coral species available in the IUCN repository (<http://www.iucnredlist.org>). Unfortunately, these are the only marine groups available in the IUCN repository among the taxonomic groups in our study.

We also evaluated the range cohesion of the marine species using estimates of species distribution models, which are less sensitive to geographical gaps in sampling effort and have been considered a good alternative to estimate complete species range (Herkt *et al.* 2017). All the modeled range maps were retrieved from AquaMaps (<https://www.aquamaps.org>), which provides standardized range maps for marine species based on environmental tolerances with respect to depth, salinity, temperature and primary production. A large part of species with range maps available in the IUCN repository also had estimated ranges in AquaMaps, except for corals (Scombridae: 98%, Elasmobranchii: 71%, Corals: 0.8%).

Our results show that range contiguity is the most common distribution pattern in nature (Supplementary Fig. 1). Only 10.7% of species have disjunct latitudinal range according to IUCN's range maps, whereas this estimate drops to only 2.7% according to species distribution models. Conversely, according to OBIS occurrence records, 57.3% have range discontinuities at 5° latitudinal resolution. It is also noteworthy that the equatorial dip in species diversity tends to decrease, or even disappear, when using range maps

(Supplementary Fig. 1). Therefore, our results suggest that the geographical distribution of most marine species tends to be contiguous across space.

Supplementary Note 2 | Bathymetric distribution of marine species records

Records from relatively shallow waters (0-2,000 m) represent 89.23% of all records analysed in our study. Considering only the records with depth information and within our range of analysis (0 – 6,500 m), euphotic and bathyal strata represent 99% of all records (euphotic: 87.97%; bathyal: 11.10%), meaning that less than 1% of information came from the deep sea, the largest habitat on Earth. The maximum sampling effort per latitudinal band for any taxon, for example, decreased two orders of magnitude toward deeper strata (euphotic: 59,288; bathyal: 14,775; abyssal: 313). Consequently, the estimates of inventory completeness were, on average, lower at the deep sea than in shallow waters, especially in the southern hemisphere (see Fig. 3f in the main text). The total number of species, and species with gaps in their latitudinal range, also decreased 4.78 and 4.74 times, respectively, from euphotic to the deep-sea stratum (euphotic: 26,538 - 8,262; bathyal: 15,828 - 4,735; abyssal: 5,551 - 1,744).

Supplementary Note 3 | Standardizing sample size and completeness

Because tropical marine biota is undersampled, a direct comparison of species richness across latitudinal bands is not appropriated. However, two strategies may be employed to overcome this problem. First, it is possible to standardize the sample size by randomly subsampling the records across latitudinal bands to the same level as observed in the tropics. We repeated this procedure 1000 times and calculated the average latitudinal variation in species richness and completeness estimate expected in a scenario of

homogeneous sampling effort across the globe. Note that we did not standardize the number of records, but the number of sampling events. Therefore, the number of records by sampling event was not altered. Polar latitudes were not subsampled because they already have sampling intensity below the tropical threshold.

Second, it is also possible to standardize the sampling coverage (interpolation; Chao & Jost 2012). Using the *iNEXT* R package (Hsieh *et al.* 2016), we reduced the sampling effort across latitudes to the lowest completeness estimate observed in the tropics (between -15° and 15°). Additionally, we extrapolated tropical species richness. It is noteworthy that such extrapolation should be extended only to a doubling of the reference sample size. Beyond that level the extrapolation could create unreliable estimates. For this reason, we did not extrapolate the species richness for all latitudes until estimated total completeness. Instead, we used the extrapolation to achieve a higher standardized sampling coverage (Chao & Jost 2012).

Our results show that standardization of sampling effort have a great impact on the latitudinal pattern in marine species richness. For most groups, when sampling effort have equal size, the number of species at those rich mid-latitudes drops to a value similar or lower than that recorded near the equator (Supplementary Fig. 2). Interestingly, if this reduced sampling effort was real, many species would not have been recorded at mid-latitudes. The reduction in sampling effort and species richness also altered the completeness estimate. However, the changes were more modest at higher latitudes, especially for sample coverage (Supplementary Fig. 4-5). This result suggests that while many tropical species are rare (few records), species at high latitudes are equally abundant and well represented, keeping completeness estimate relatively high even under reduced sampling effort. The latitudinal difference in completeness estimate despite standardized sample size also reinforces that a homogeneous sampling effort could produce a heterogeneous latitudinal gradient in

inventory completeness. Likewise, heterogeneous sampling effort could also produce homogeneous inventory completeness. Therefore, species rich areas require more sampling effort to reach high completeness, especially if most species are rare or spatially restricted.

The results of coverage-based standardization also showed that species richness tends to be higher near the equator, for both interpolation and extrapolation scenarios (Supplementary Fig. 17-18). Indeed, for some groups, the extrapolated species richness was even higher in the tropical dip than around it (see for example Ophiuroidea, Amphipoda, and Porifera in Supplementary Fig. 17). Interestingly, the extrapolation also revealed that doubling the sample size and increasing the completeness estimate had greater impact on tropical species richness, but almost did not affect the diversity at higher latitudes. Our results suggest that species richness should increase rapidly with additional sampling effort in the tropics, but substantial effort would be necessary to reach the same level of completeness as currently existing at higher latitudes.

Supplementary Note 4 | OBIS representativeness and inventory completeness estimate

To evaluate if data retrieved from OBIS is representative of current knowledge of global patterns of marine biodiversity, we compiled ophiuroids data from additional nine data repositories worldwide. All retrieved data were submitted to the same quality control procedures described in the main text. OBIS records represented 57.81% of all valid records in the full dataset (Supplementary Table 1). We then calculated independently for each dataset the number of observed species (richness observed), number of absent species (spatial gaps), number of sampling events, and estimated inventory completeness for each 5° latitudinal bands.

Following Stropp *et al.* (2016), we employed three methods to estimate inventory completeness. First, we applied Sousa-Baena *et al.* (2014) estimation by generating tables of unique combinations of species name, coordinates and date of collection. For each latitudinal band we then calculated the total number of observed species (S_{obs}), and the number of observed species in only one (a) or two (b) sampling events. Sampling event is a unique combination of latitude, longitude and date of collection. Sousa-Baena *et al.* (2014) estimation can be calculated by:

$$C_i = \frac{S_{obs\ i}}{S_{obs\ i} + (a_i^2 / 2b_i)}$$

where C_i is the inventory completeness for the latitudinal band i . Completeness estimate cannot be calculated when the parameter b is not found.

The second method to estimate inventory completeness uses sample coverage (Chao & Jost 2012). For each latitudinal band we quantified the total number of records (n), and number of species observed in only one (f_1) or two (f_2) sampling events. Sample coverage can be calculated by:

$$K_i = 1 - \frac{f_1}{n} \left[\frac{(n-1)f_1}{(n-1)f_1 + 2f_2} \right]$$

where K_i is the estimated inventory completeness of latitudinal band i . Both C and K ranges from zero to one, with one indicating a complete inventory.

Finally, we also estimated inventory completeness by the curvilinearity of smoothed species accumulation curves (SACs; Tittensor *et al.* 2010; Yang *et al.* 2013). Smoothed SACs were calculated by the ‘exact’ method (function ‘specaccum’ in the R package *vegan*; Oksanen *et al.* 2018). The average slope of the last 10% of SACs obtained for each latitudinal band reflects the degree of curvilinearity, and was used to estimate the inventory completeness (r_i). A flat slope indicates saturation in the sampling (high inventory completeness), but produces r_i value close to zero. We calculated the 1-completeness

estimate r_i to convert the values to a normalized scale from zero to one, in which one indicate high inventory completeness (Stropp *et al.* 2016).

We used Pearson's product-moment correlation to evaluate the congruence of latitudinal gradients in species richness, spatial gaps and sampling effort estimated using only OBIS and the full dataset (OBIS + additional datasets). Our results revealed a strong correspondence between the two datasets (Supplementary Table 2), suggesting that additional datasets are not substantially different from OBIS (Supplementary Fig. 1). Estimates of inventory completeness are also not affected by the dataset, suggesting that OBIS can be used in isolation to estimate the latitudinal pattern in inventory completeness (Supplementary Table 2).

Estimates using Sousa-Baena *et al.* (2014) method were not so correlated between the two datasets when compared to the other two completeness estimates. In addition, the monotonic relationship between completeness estimate and number of records seems more unstable under Sousa-Baena *et al.* (2014) method (Supplementary Fig. 22), suggesting that this method is more susceptible to artefactual values of C in latitudinal bands with small number of sampling records (Sousa-Baena *et al.* 2014; Stropp *et al.* 2016). Because many latitudinal bands have few sampling records of some taxonomic groups, we abandoned Sousa-Baena *et al.* (2014) method in all further analyses.

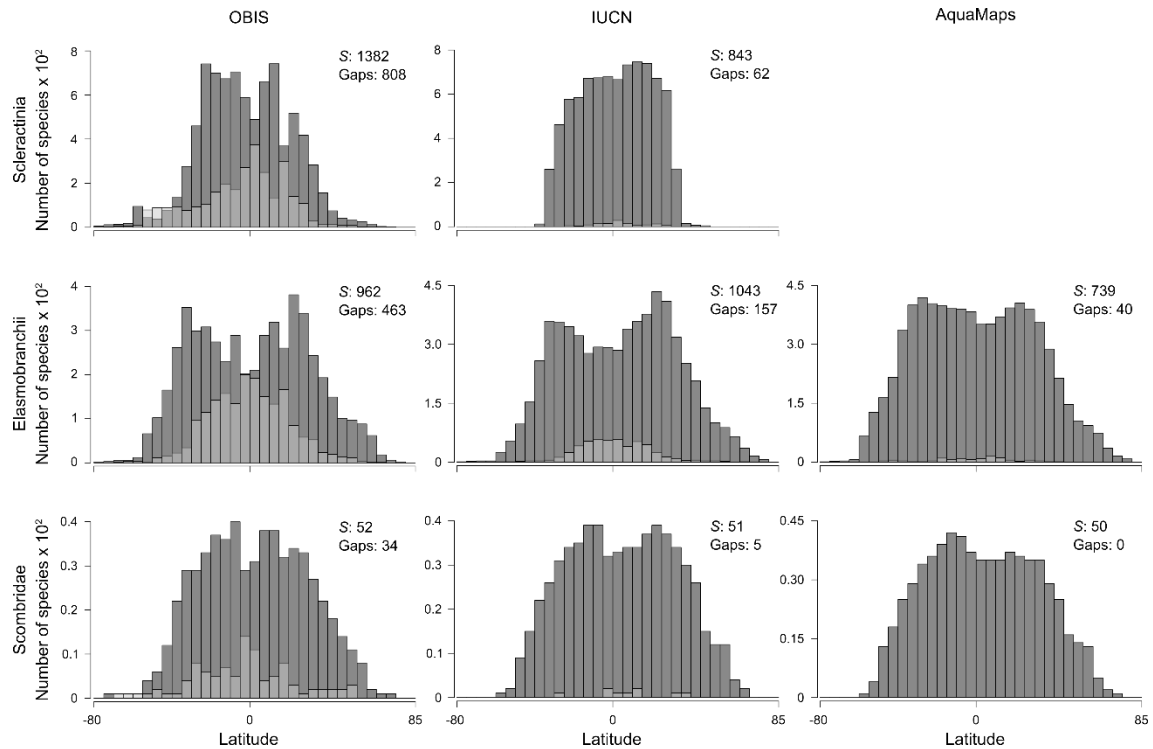
Supplementary Note 5 | Data cleaning procedures

Over half of our initial dataset retrieved from OBIS (and additional repositories for Ophiuroidea) was eliminated during the data cleaning processes (65.82%; Supplementary Table 3). One third (33.81%) of excluded records were duplicates, and a quarter (26.51%) lacked species-level identification. Although exclusion of duplicates has no effect on presence of latitudinal gaps in species distribution, the effect of excluding records with

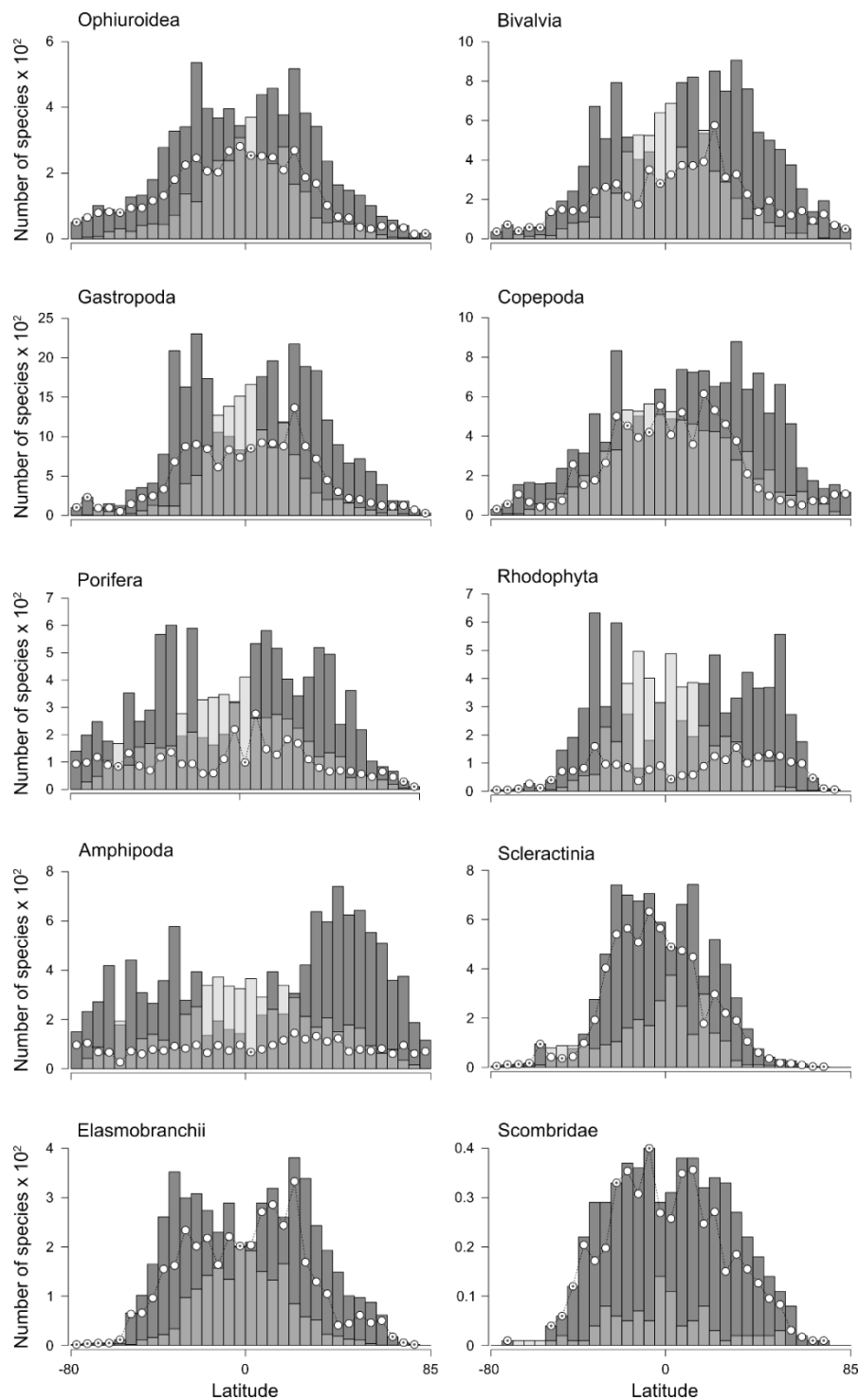
missing species identification requires further investigation. In fact, the predominance of such records in tropical waters could indicate that tropical species absence may be caused by lack of taxonomic expertise, instead of low sampling effort. To further investigate such possibility, we studied the latitudinal distribution of records lacking species-level identification. The data used in this analysis was retrieved on May 30 (2018), and is not identical to the dataset used in our main analyses, retrieved months earlier. Comparison between the two datasets shows an increase of 7.1% in the number of total records, but an equivalent proportion of records missing species-level identification (29.97%). Most records with unidentified species are located in well-sampled mid-latitudes (Supplementary Fig. 20). While equatorial data (between -5° and 5°) comprise only 2.49% of all records without identification, 29.4% of records lacking species-level identification are located between 50° and 60° . Thus, exclusion of records lacking species-level identification during the data cleaning process does not seem to cause spatial gaps in species' ranges.

Another potential cause of spatial gaps at low latitudes could be the removal of records with both coordinates equal to zero. Such records were eliminated because 0-0 location is probably auto-filled by computers when coordinate fields are left blank (Vandepitte *et al.* 2015). Because there are not many records at the 0-0 location (Supplementary Table 4), the removal of these records should have minimal impact for the latitudinal pattern of species absence. However, the 0-0 location is an actual marine location in the Atlantic Ocean (Gulf of Guinea) and some of the excluded records at this location could potentially represent accurate records. Thus, we identified all species with records at the 0-0 location, and then quantified the proportion of species with range gaps. We found that less than 1% of the species with spatial gaps had any excluded 0-0 location record (Supplementary Table 4). In addition, most of the records for benthic taxa indicate an unrealistic shallow depth. While the 0-0 location has an actual depth around 4,900 meters,

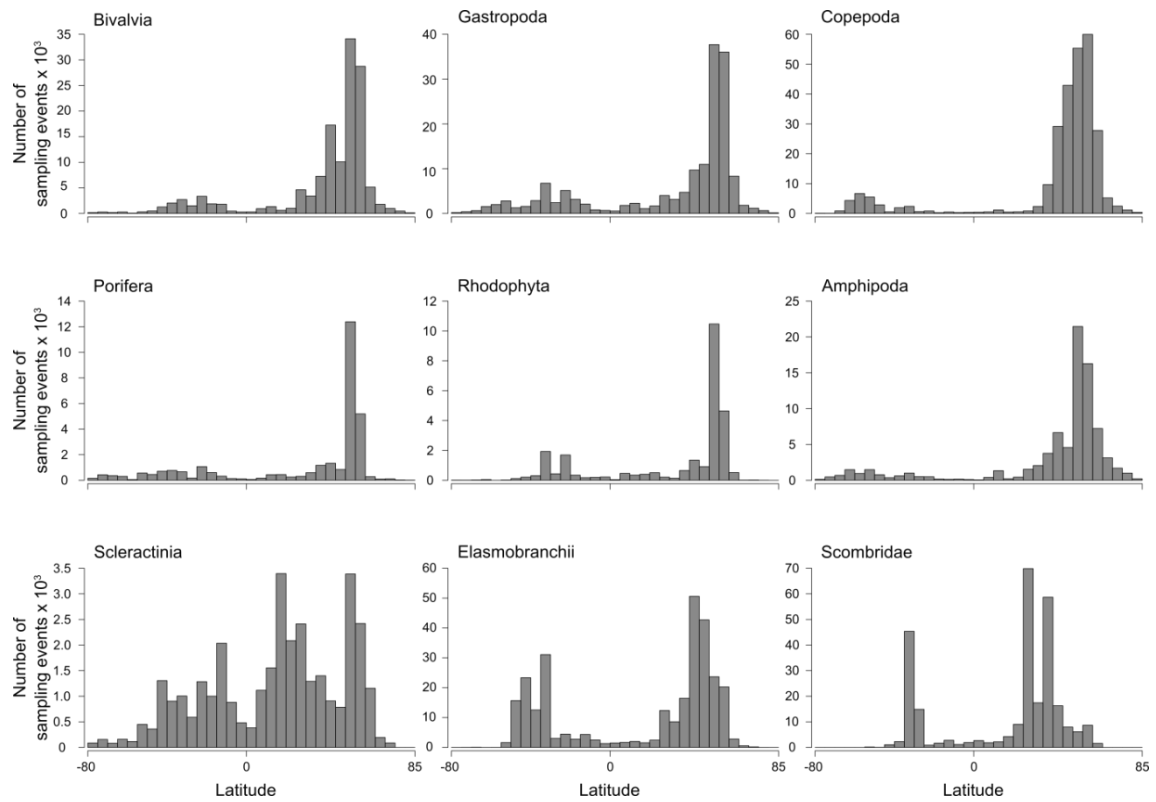
most of the 0-0 records for benthic taxa indicate a depth between 8 and 14 meters, clearly indicating auto-filled and wrong geographical coordinates.



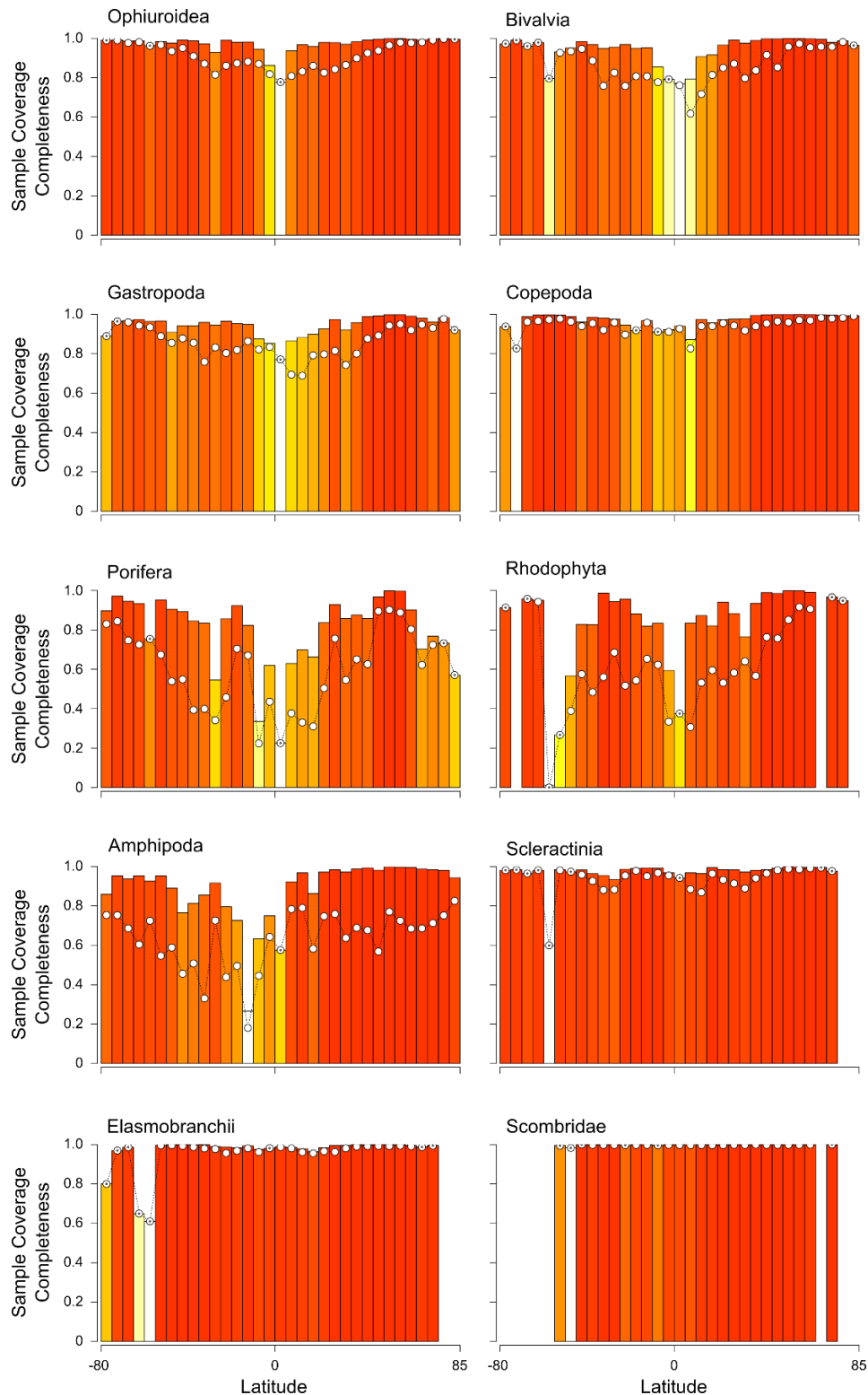
Supplementary Figure 1 | Latitudinal pattern in species richness according to different data sources. Dark grey bars represent number of observed species; light grey bars represent number of missing species (spatial gaps); colors are mixed where histograms overlap. S: total species richness; Gaps: number of species with any spatial gap. Notice that tropical peak in missing species coincides with the tropical dip in species richness. Range-based richness patterns (IUCN and AquaMaps) are not as bimodal as occurrence-based richness patterns (OBIS).



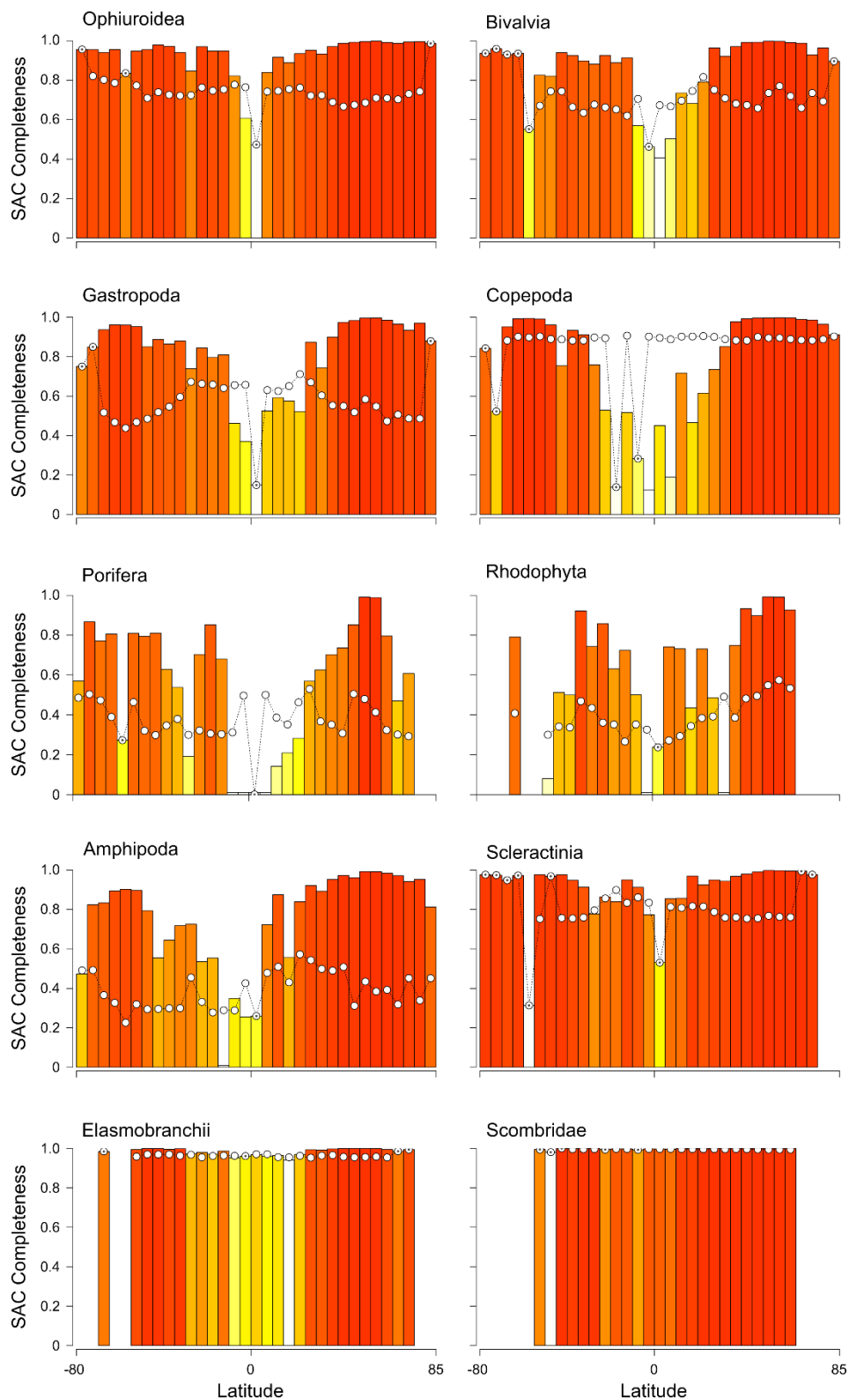
Supplementary Figure 2 | Latitudinal pattern in species richness of different taxonomic groups. Dark grey bars represent number of observed species; light grey bars represent number of missing species (spatial gaps); colors are mixed where histograms overlap. Circles represent mean ($\pm$ s.d.) expected species richness under equally low sampling effort (see details in Supplementary Note 3). Filled circles indicate latitudes where sampling effort is already low and was not reduced. Notice that tropical peak in missing species coincides with the tropical dip in species richness.



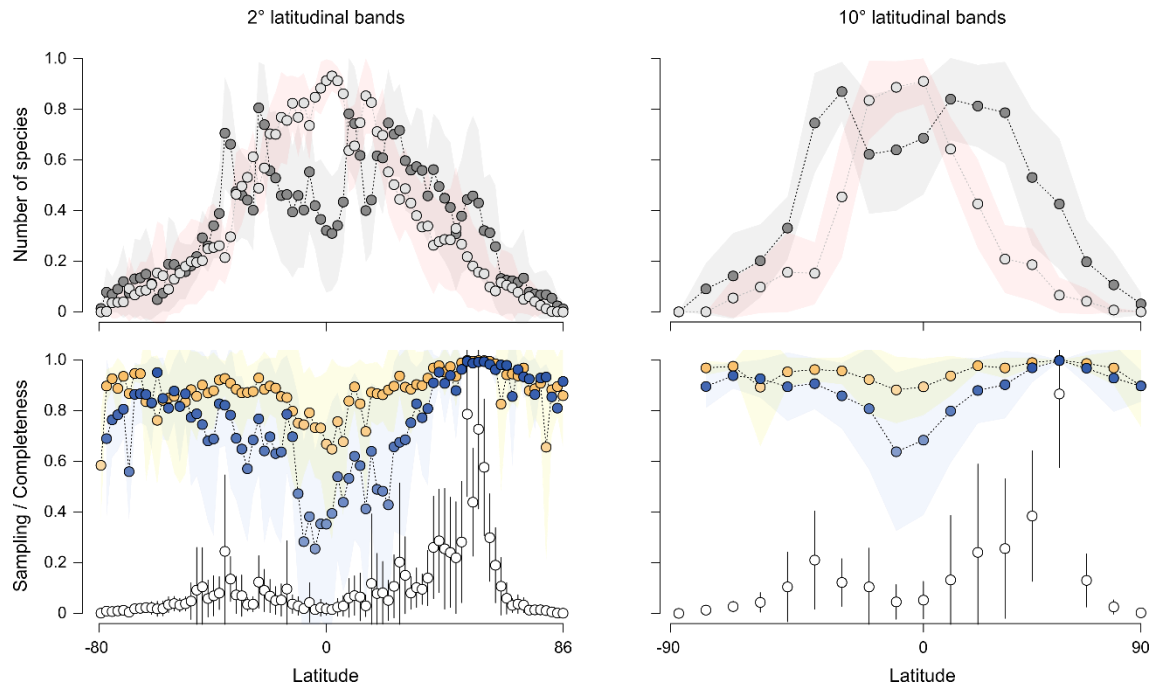
Supplementary Figure 3 | Latitudinal pattern in sampling events of each taxonomic group. See Supplementary Fig. 19 for Ophiuroidea graph.



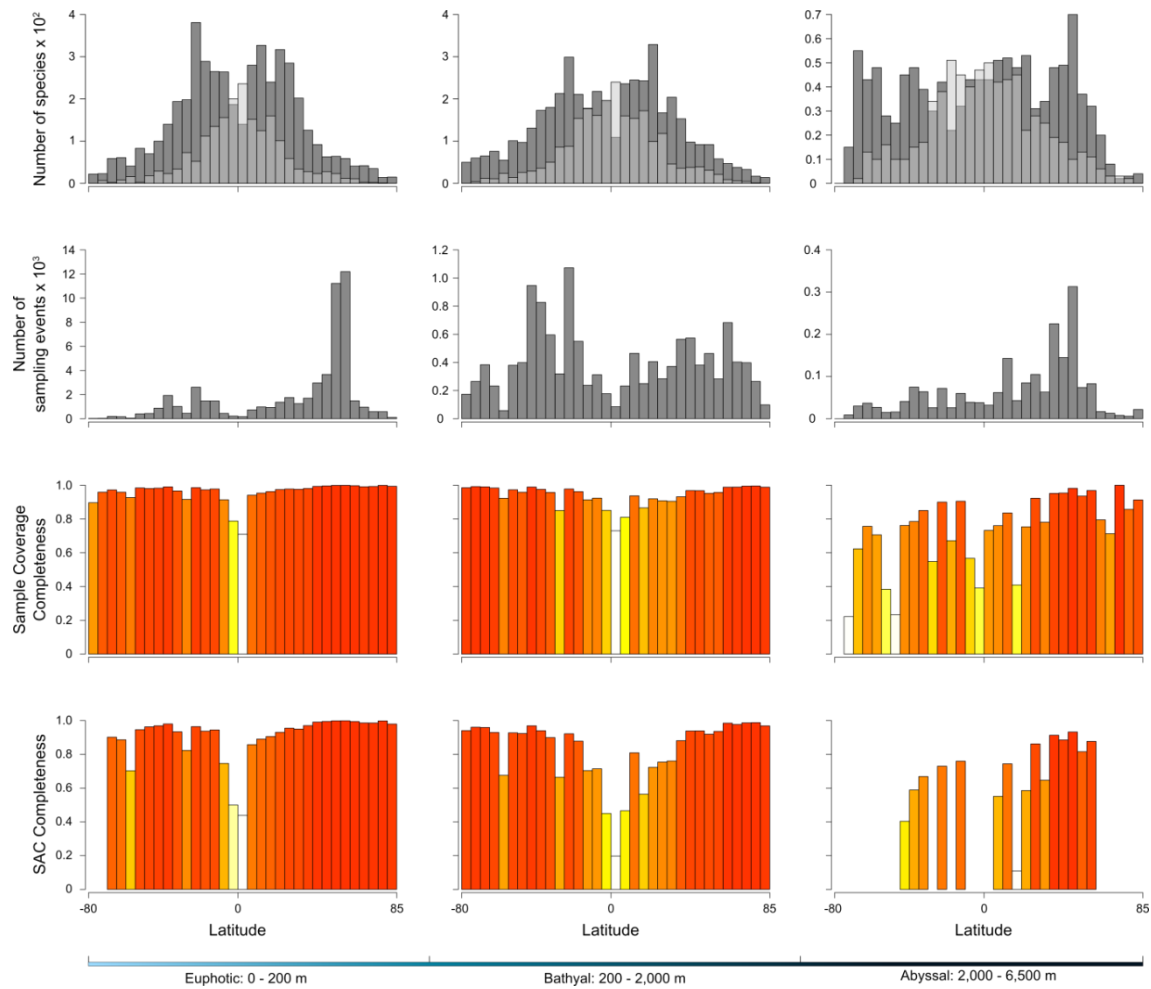
Supplementary Figure 4 | Inventory completeness of each taxonomic group estimated by the sample coverage method. Red indicates maximum and white indicates minimum estimated completeness level. Circles represent mean ($\pm$ s.d.) expected inventory completeness under a scenario of homogeneous sampling effort across the globe (see details in Supplementary Note 3). Filled circles indicate poorly sampled latitudinal bands, in which subsampling was not applied.



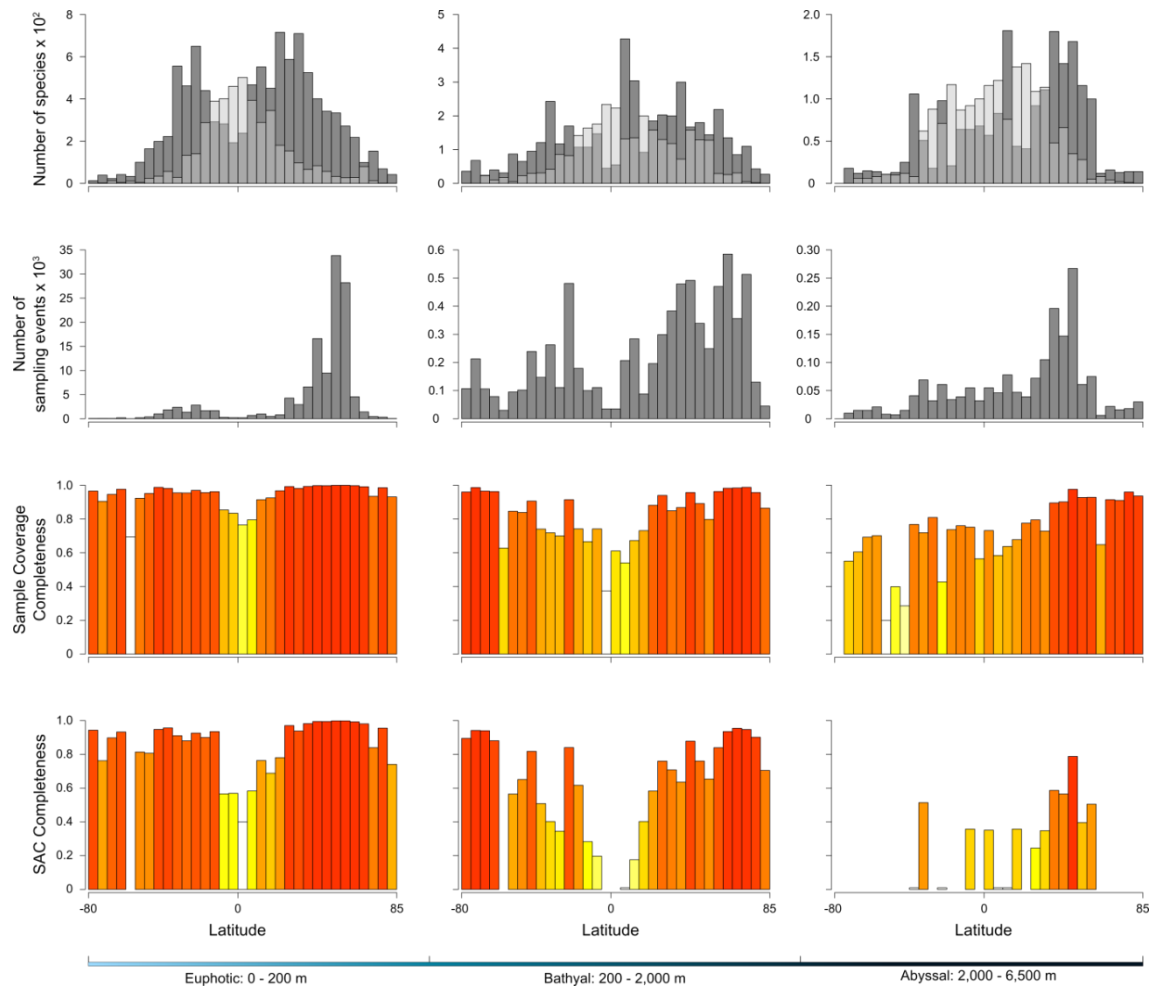
Supplementary Figure 5 | Inventory completeness of each taxonomic group estimated by the species accumulation curve method. See caption of Supplementary Fig. 4.



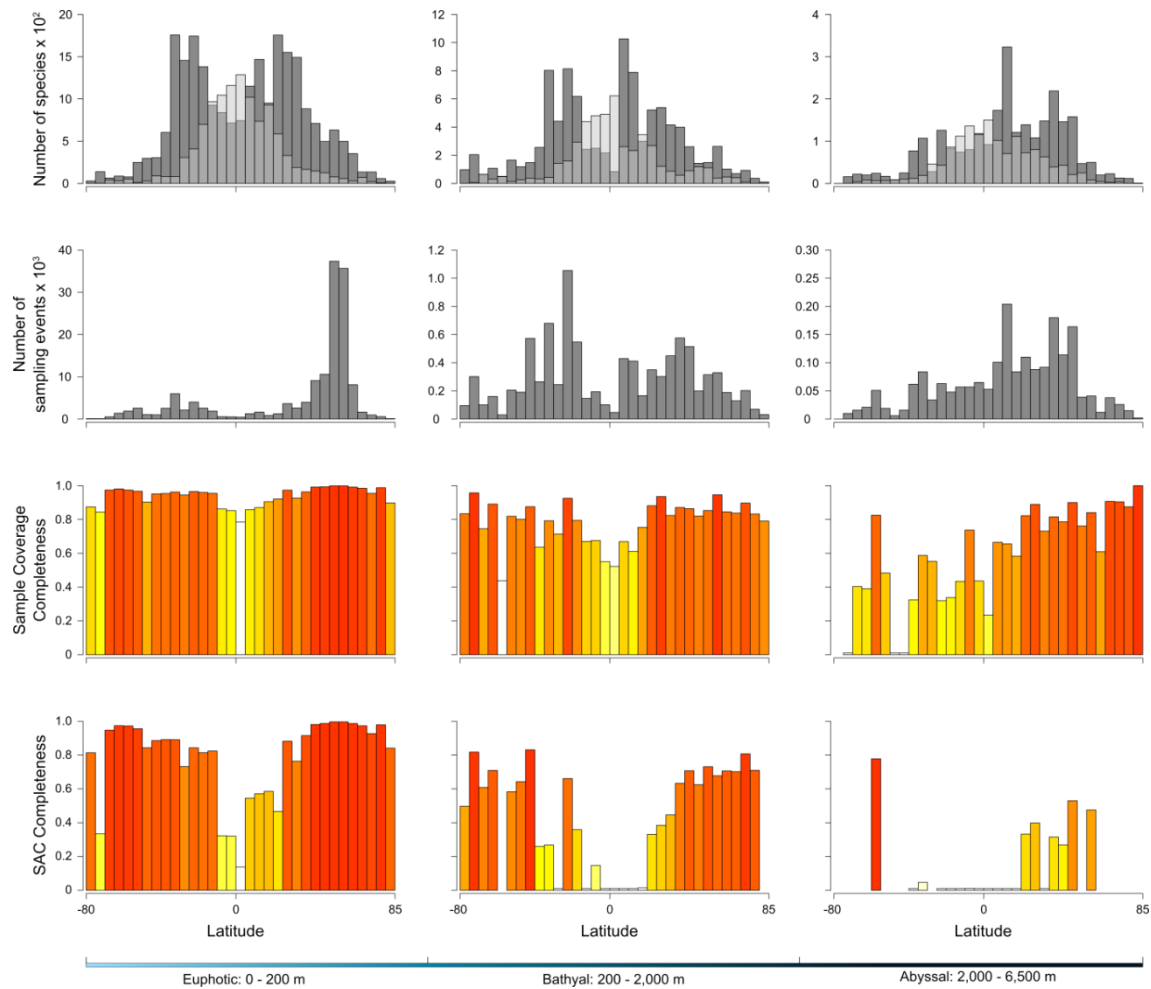
Supplementary Figure 6 | Latitudinal patterns in species richness, spatial gaps and sampling efficiency for marine taxa under different spatial resolutions. Circles represent mean values based on results of ten taxonomic groups. Shaded area and vertical bars represent standard deviation (s.d.). Data were standardized by the maximum observed value to range between 0 and 1 (except those which already vary on this scale). Top: Mean ($\pm$ s.d.) latitudinal variation in number of observed species (dark grey circles) and number of missing species (light grey circles). Bottom: Mean ($\pm$ s.d.) latitudinal variation in sampling effort (number of unique sampling events; open circles) and two estimates of inventory completeness (sample coverage: orange circles; species accumulation curve: dark blue circles).



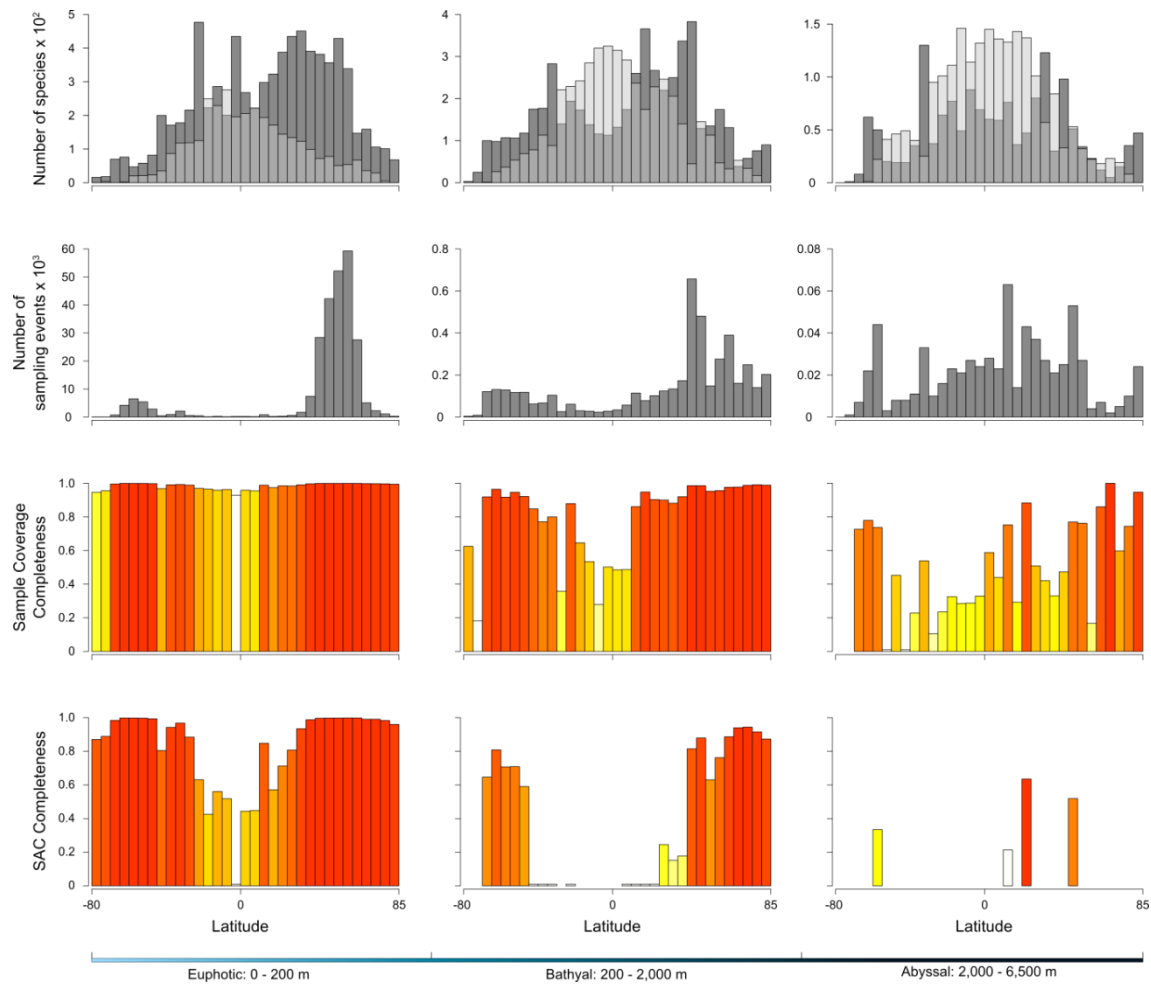
Supplementary Figure 7 | Latitudinal patterns in species richness (observed and missing), number of sampling events, and estimates of inventory completeness for Ophiuroidea at three different depth strata (euphotic, bathyal and abyssal). For species richness graphs, dark grey bars represent number of observed species; light grey bars represent number of missing species (spatial gaps); colors are mixed where histograms overlap. For inventory completeness, red indicates maximum and white indicates minimum level of completeness in the plot.



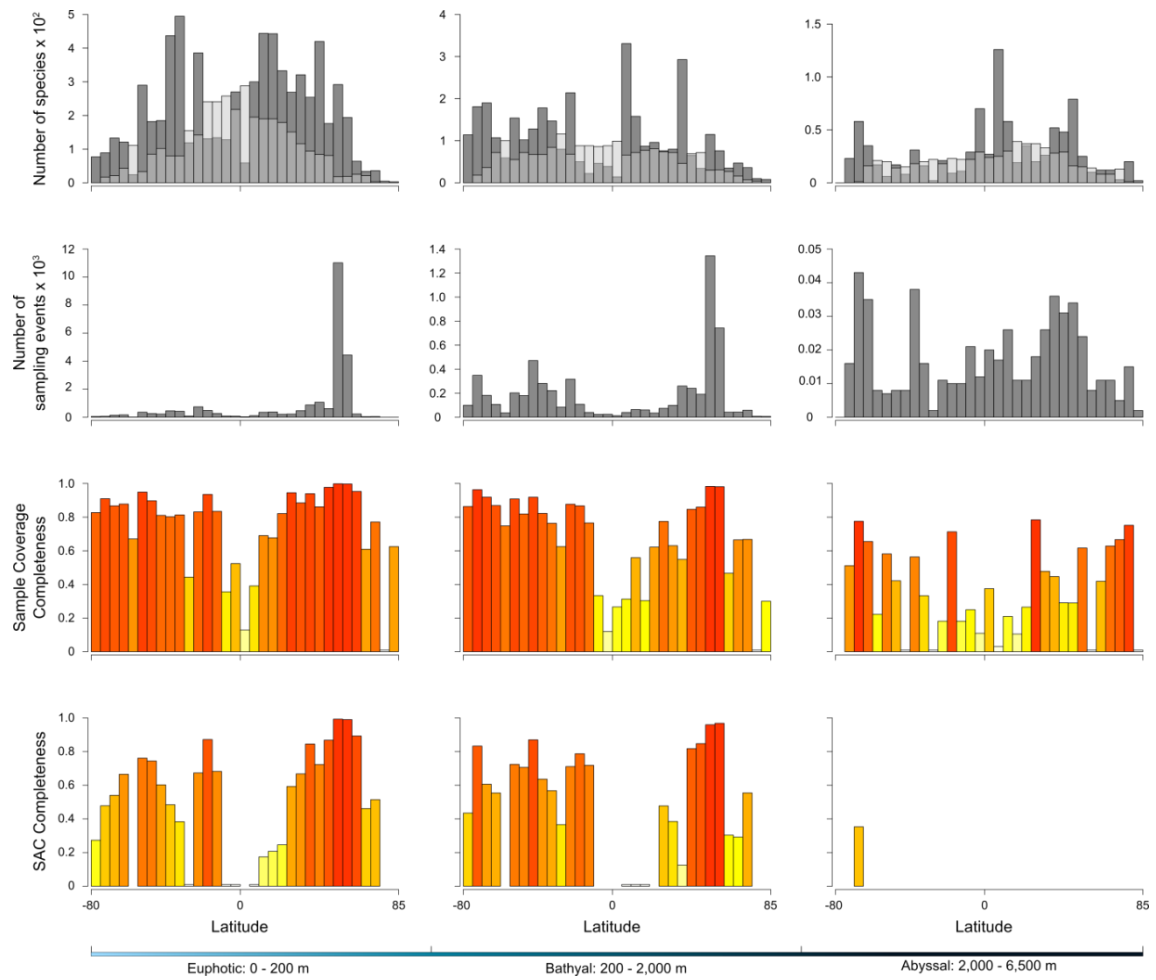
Supplementary Figure 8 | Latitudinal patterns in species richness (observed and missing), number of sampling events, and estimates of inventory completeness for *Bivalvia* at three different depth strata (euphotic, bathyal and abyssal). See caption of Supplementary Figure 7.



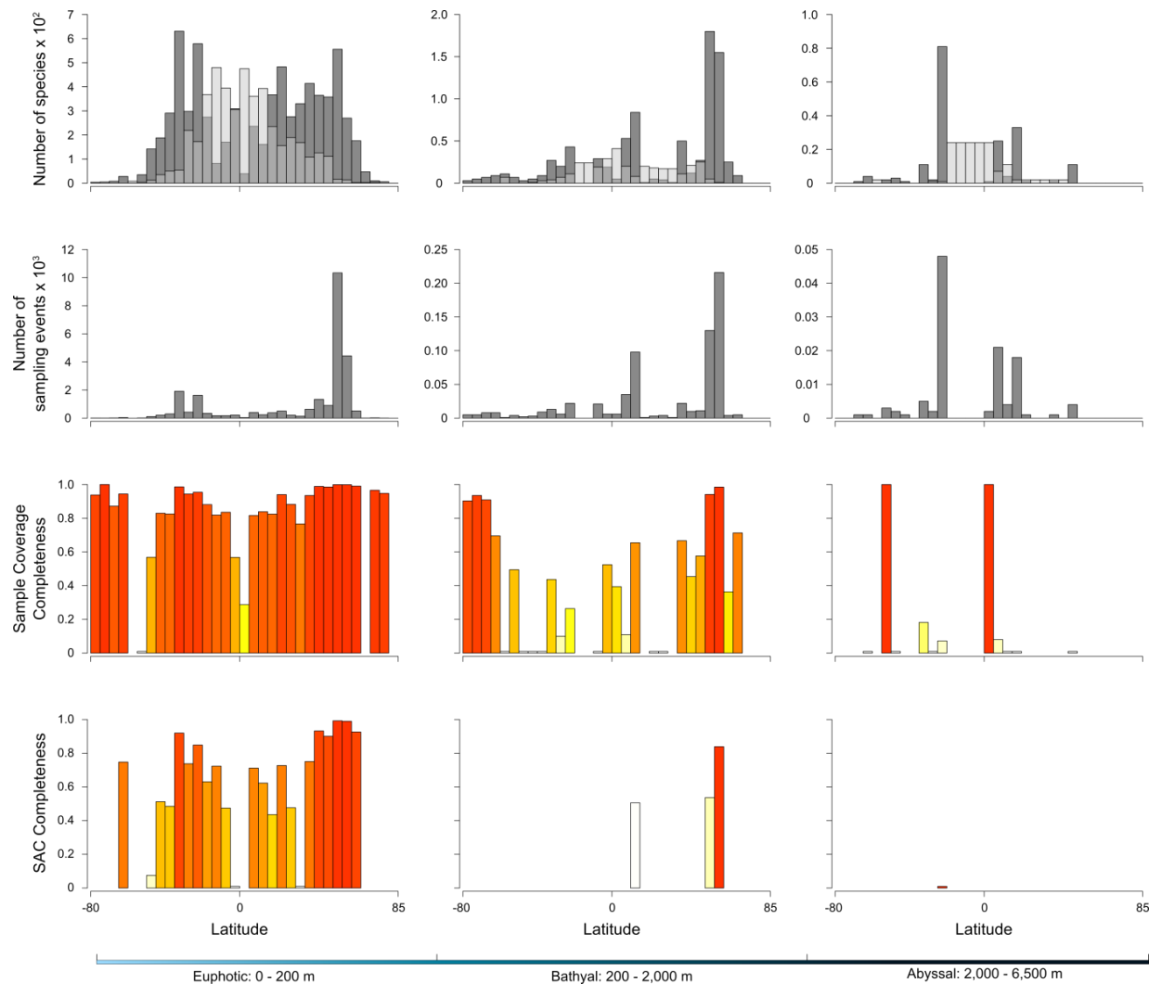
Supplementary Figure 9 | Latitudinal patterns in species richness (observed and missing), number of sampling events, and estimates of inventory completeness for Gastropoda at three different depth strata (euphotic, bathyal and abyssal). See caption of Supplementary Figure 7.



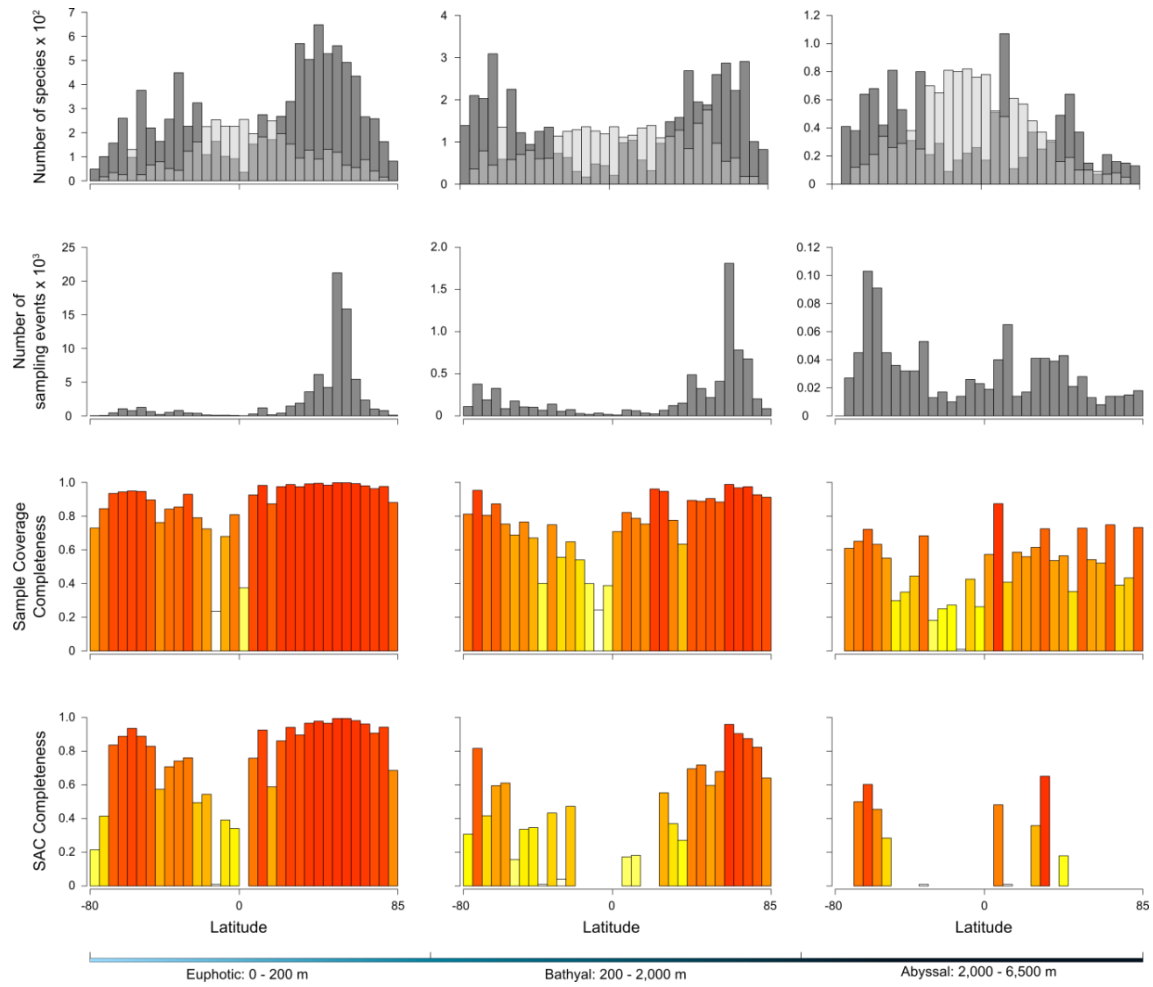
Supplementary Figure 10 | Latitudinal patterns in species richness (observed and missing), number of sampling events, and estimates of inventory completeness for Copepoda at three different depth strata (euphotic, bathyal and abyssal). See caption of Supplementary Figure 7.



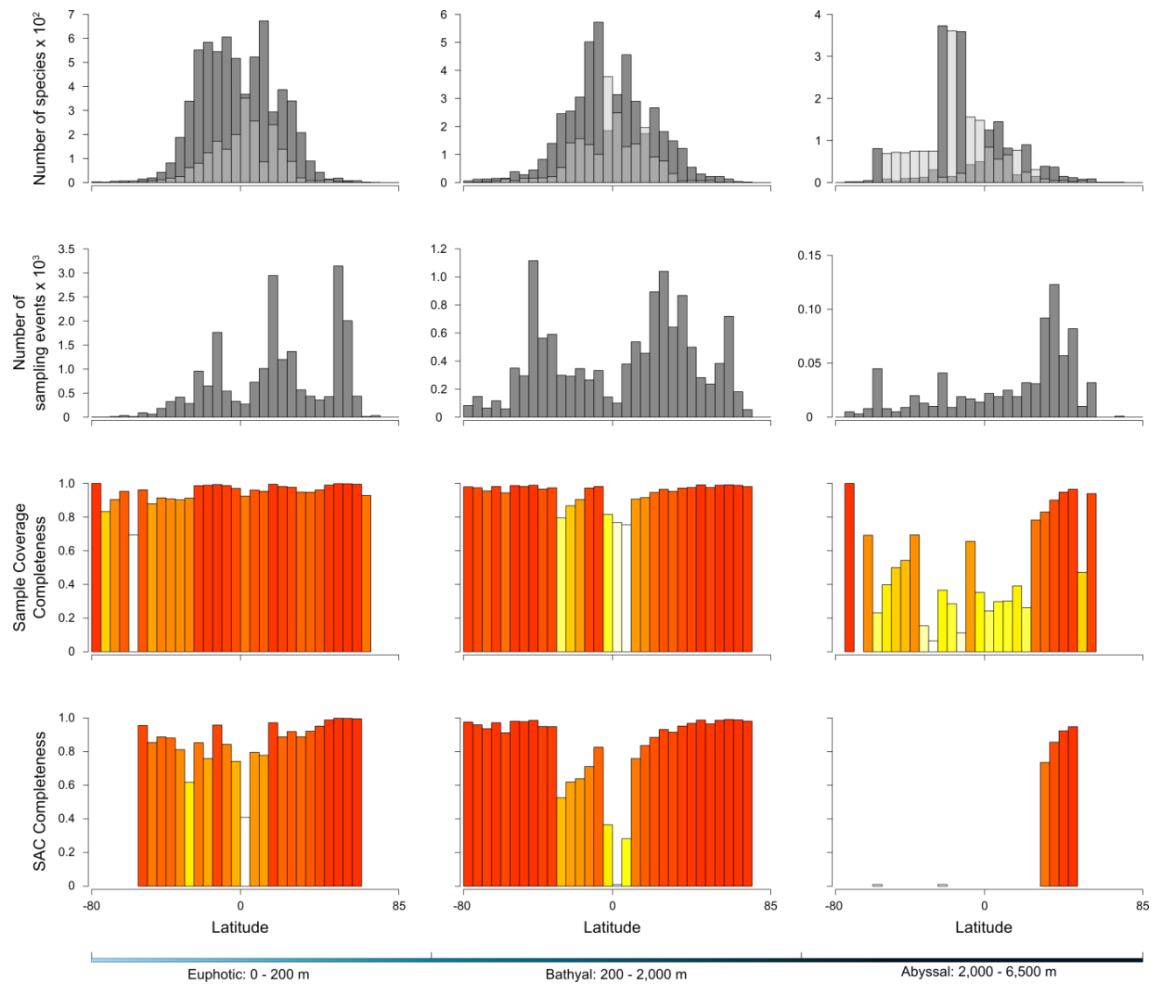
Supplementary Figure 11 | Latitudinal patterns in species richness (observed and missing), number of sampling events, and estimates of inventory completeness for Porifera at three different depth strata (euphotic, bathyal and abyssal). See caption of Supplementary Figure 7.



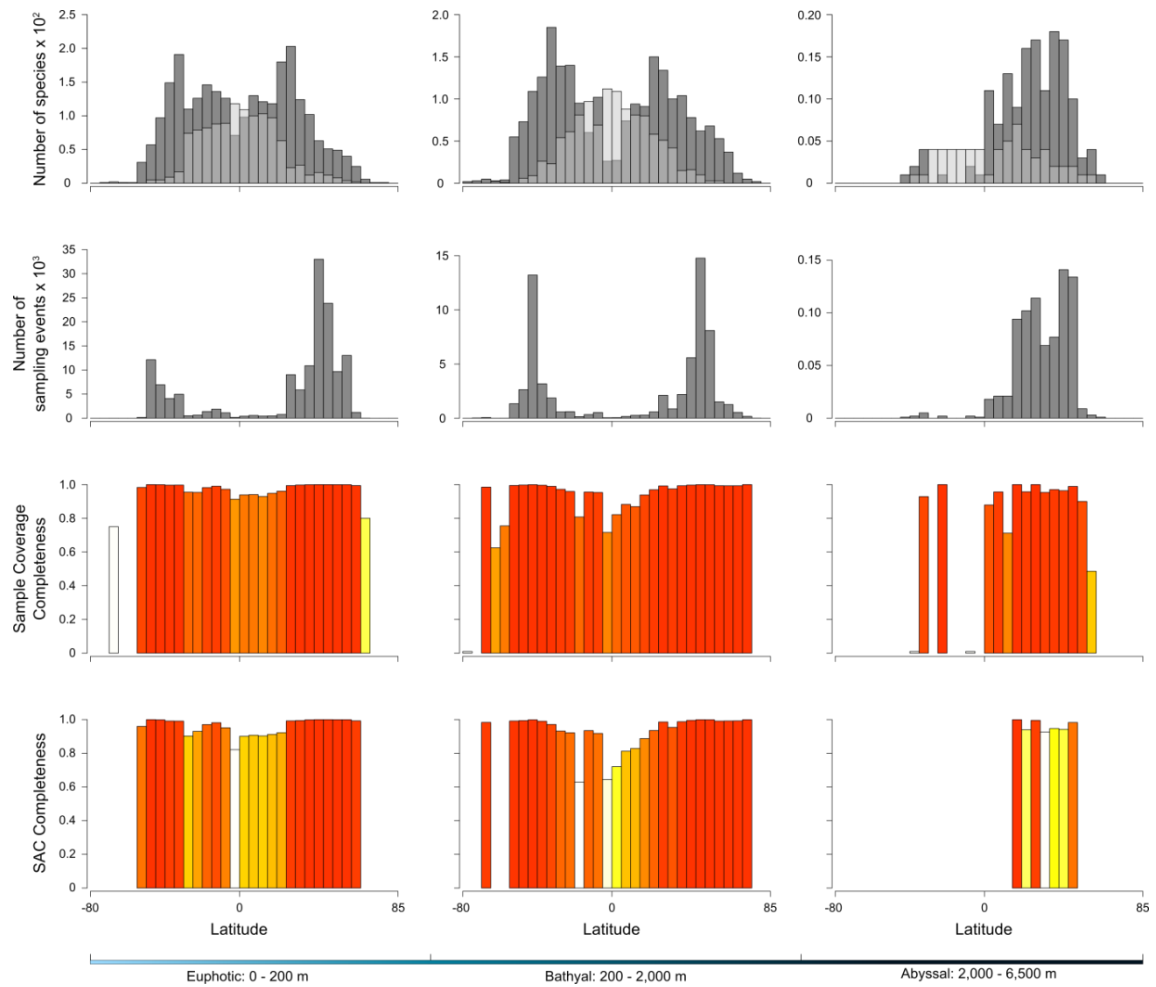
Supplementary Figure 12 | Latitudinal patterns in species richness (observed and missing), number of sampling events, and estimates of inventory completeness for Rhodophyta at three different depth strata (euphotic, bathyal and abyssal). See caption of Supplementary Figure 7.



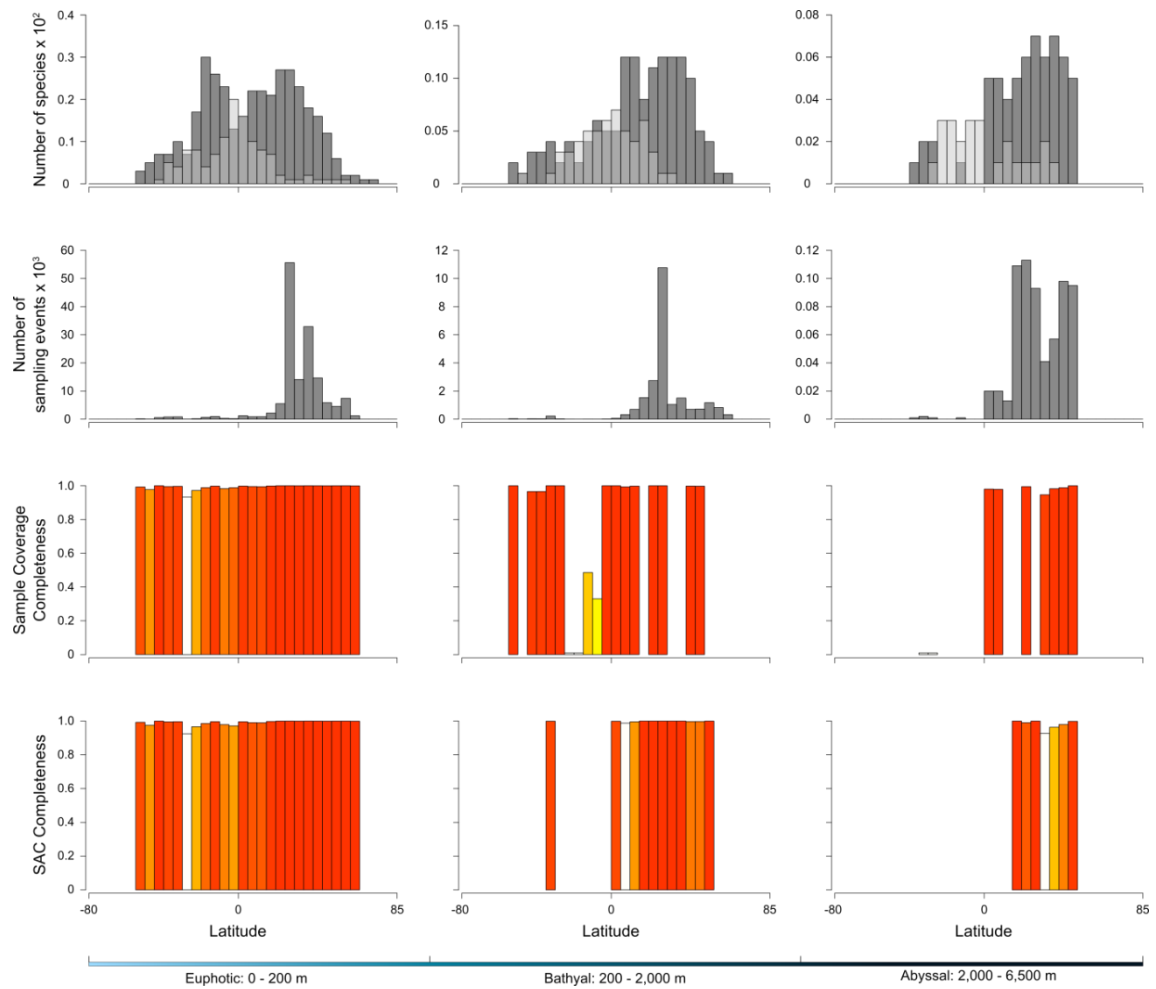
Supplementary Figure 13 | Latitudinal patterns in species richness (observed and missing), number of sampling events, and estimates of inventory completeness for Amphipoda at three different depth strata (euphotic, bathyal and abyssal). See caption of Supplementary Figure 7.



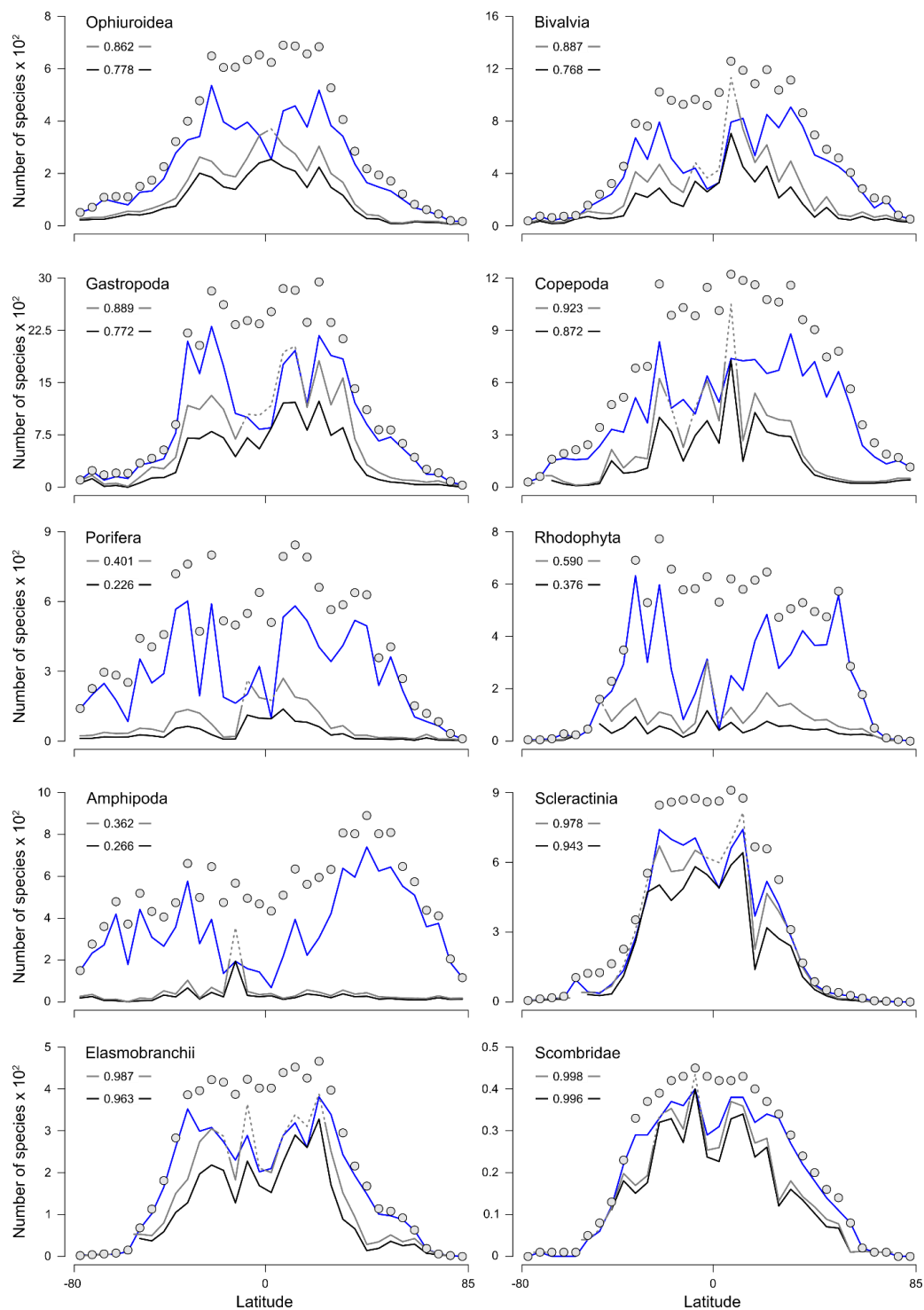
Supplementary Figure 14 | Latitudinal patterns in species richness (observed and missing), number of sampling events, and estimates of inventory completeness for Scleractinia at three different depth strata (euphotic, bathyal and abyssal). See caption of Supplementary Figure 7.



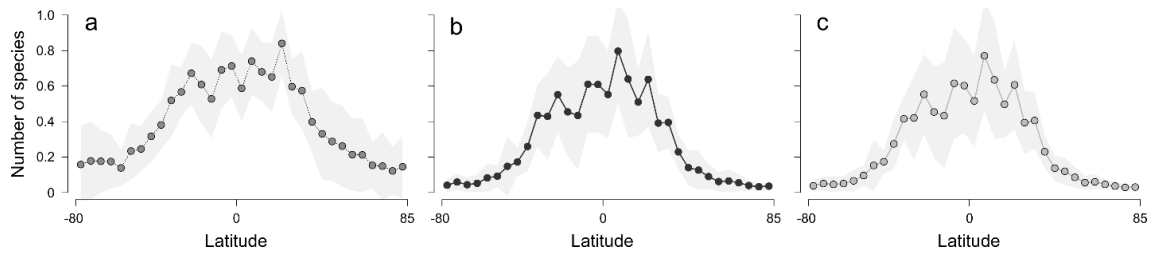
Supplementary Figure 15 | Latitudinal patterns in species richness (observed and missing), number of sampling events, and estimates of inventory completeness for Elasmobranchii at three different depth strata (euphotic, bathyal and abyssal). See caption of Supplementary Figure 7.



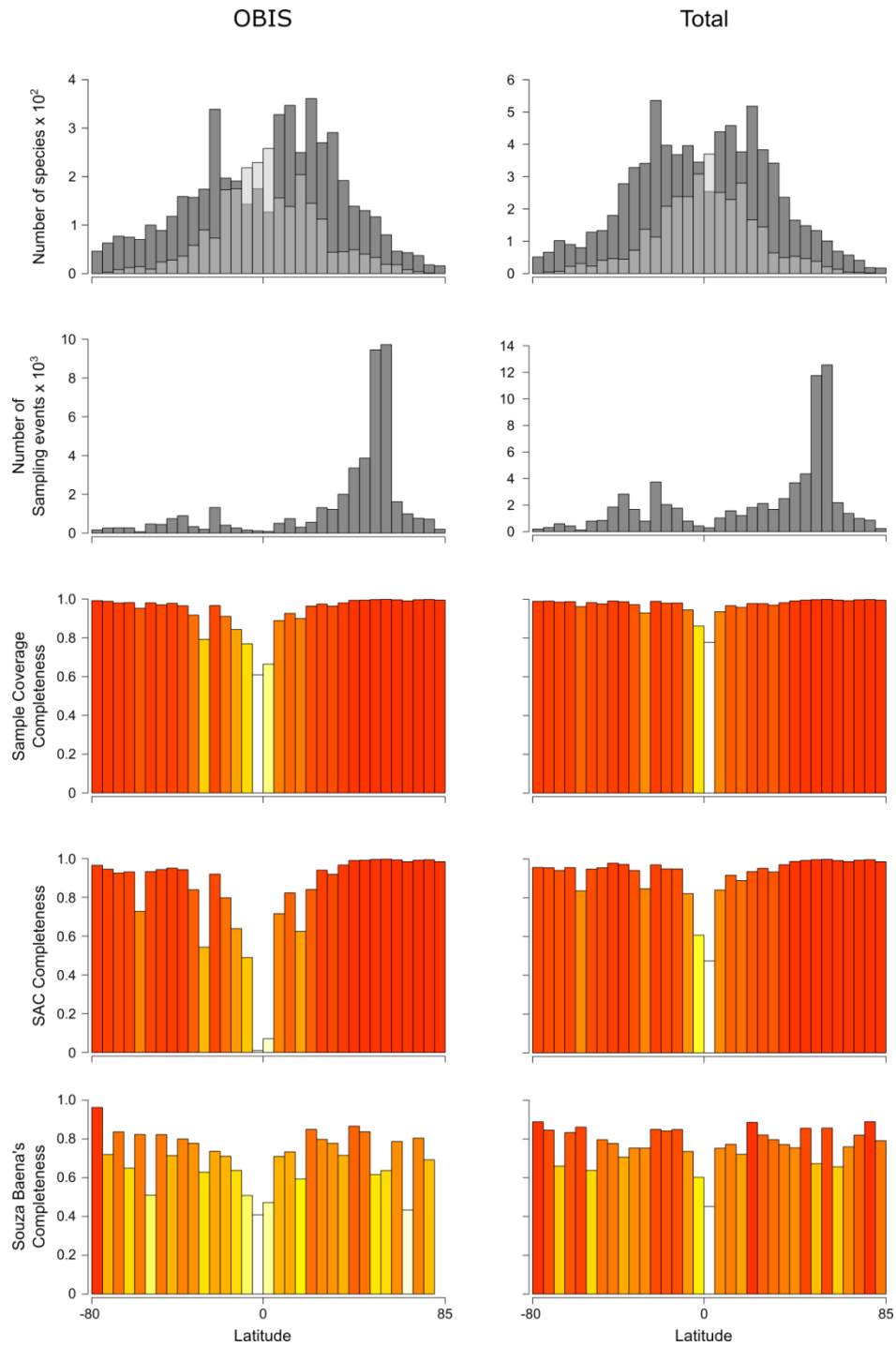
Supplementary Figure 16 | Latitudinal patterns in species richness (observed and missing), number of sampling events, and estimates of inventory completeness for Scombridae at three different depth strata (euphotic, bathyal and abyssal). See caption of Supplementary Figure 7.



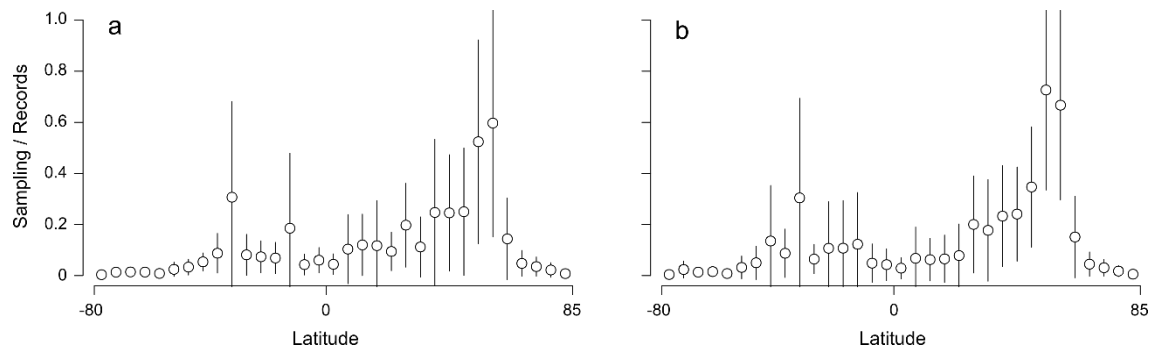
Supplementary Figure 17 | Latitudinal pattern in species richness of each taxonomic group under standardized sample coverage. Blue lines represent observed species richness; black lines represent expected species richness using only interpolation; grey lines represent expected species richness using interpolation (solid) and extrapolation (dotted). Error bars: 95% confidence interval. Numbers associated with lines (legend) indicate completeness level used to interpolate/extrapolate species richness. Grey circles represent expected species richness after accounting for spatial gaps in species distributions.



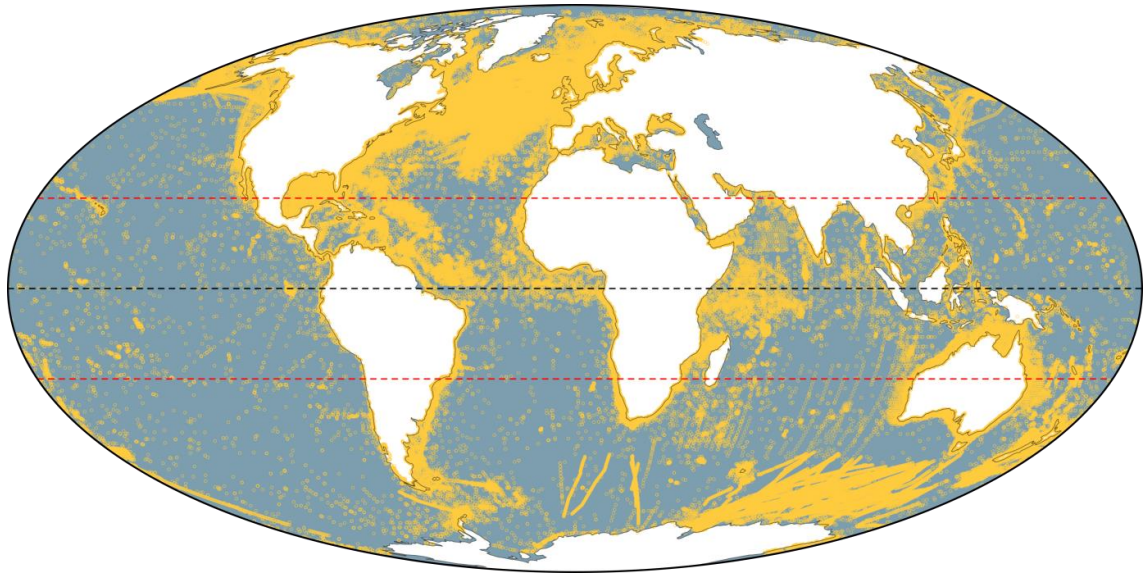
Supplementary Figure 18 | Latitudinal pattern in relative species richness under standardized sample size and coverage. Circles and shading represent mean and standard deviation of relative species richness of ten taxonomic groups. Data were standardized by the maximum observed value to range between 0 and 1. **(a)** Equal sample size. **(b)** Equal sampling coverage using only interpolation. **(c)** Equal sampling coverage using interpolation and extrapolation.



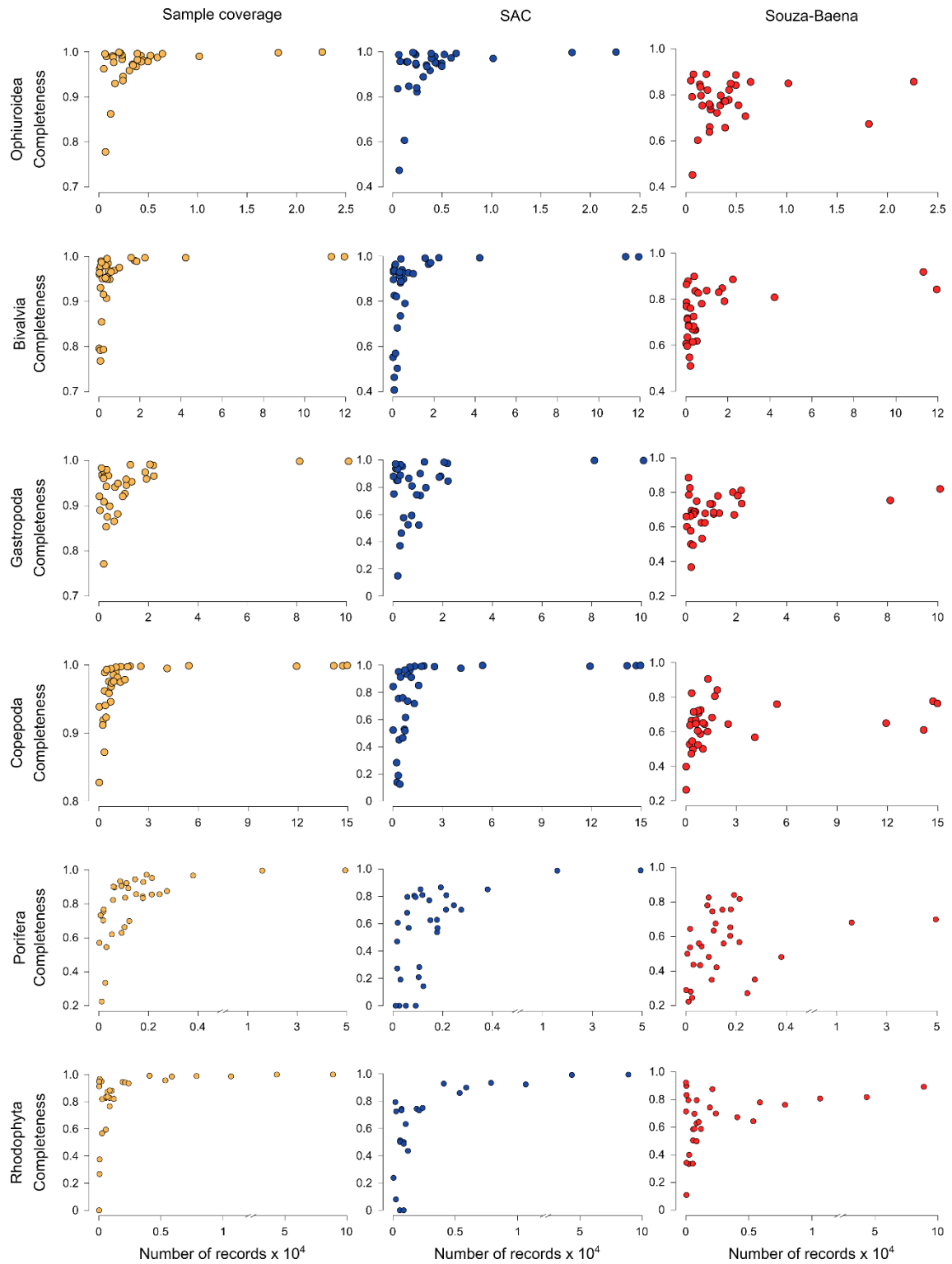
Supplementary Figure 19 | Comparison between data from OBIS and from OBIS + additional nine datasets. Latitudinal patterns in species richness (observed: dark grey; missing: light grey), number of sampling events, and estimates of inventory completeness for Ophiuroidea considering only data from OBIS (left) and from additional nine data repositories (right). See caption of Supplementary Figure 7.



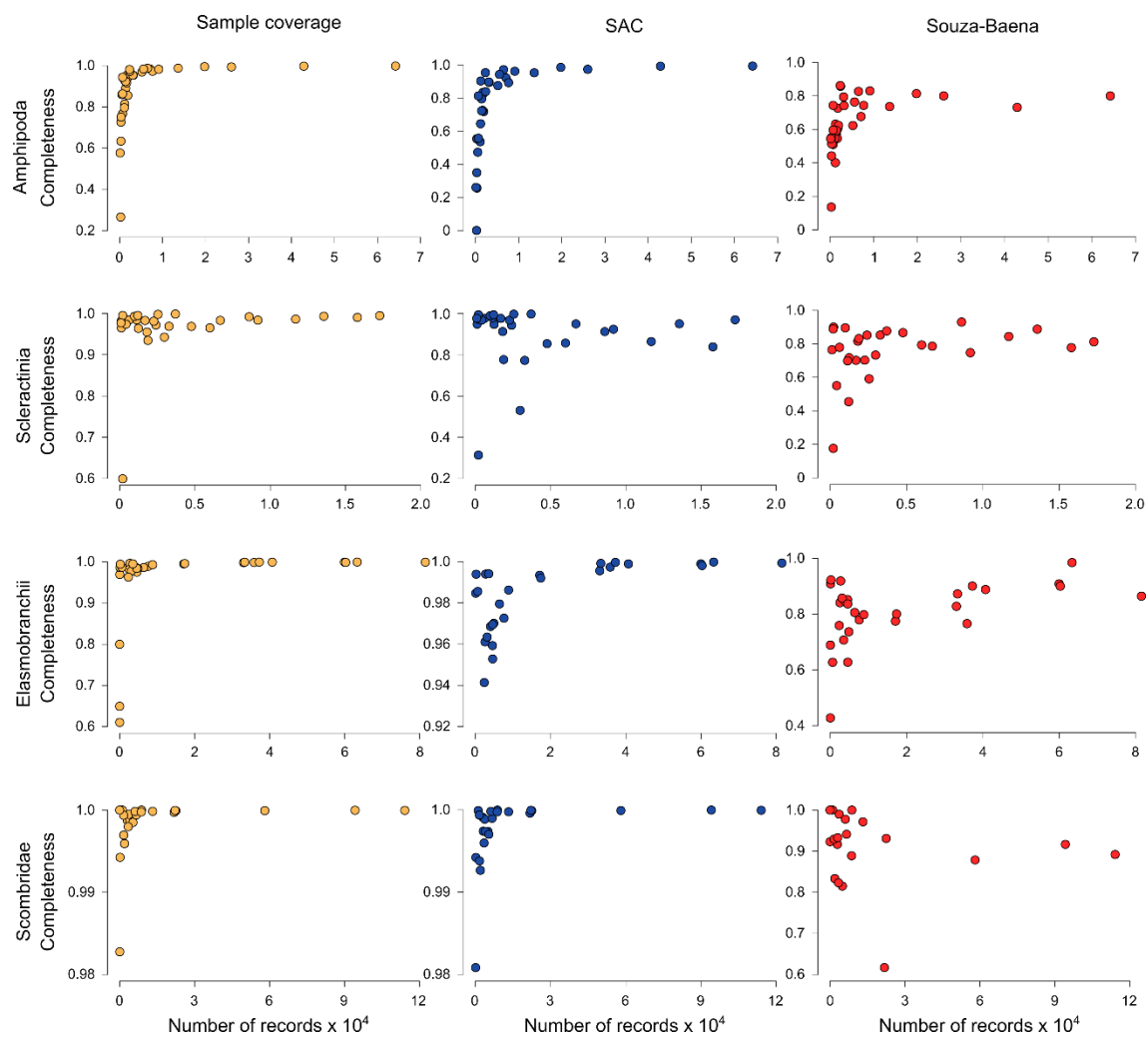
Supplementary Figure 20 | Latitudinal patterns in number of records and sampling events lacking species-level identification. Circles and vertical bars represent mean and standard deviation of ten taxonomic groups. Data were standardized by the maximum observed value to range between 0 and 1. **(a)** Total number of records. **(b)** Number of unique sampling events.



Supplementary Figure 21 | Global distribution of all sampling events used in this study. Land map used in this figure is in public domain, freely available from Natural Earth. See <https://www.naturalearthdata.com/about/terms-of-use/> for more details.



Supplementary Figure 22 | Relationship between estimates of inventory completeness and number of unique records for each taxonomic group. Points are different latitudinal bands. Notice the break in x-axis of Porifera and Rhodophyta plots. Figure continues to next page.



Supplementary Figure 22 | Continued.

Supplementary Table 1 | Number of records (total and valid) and relative proportion (%) of Ophiuroidea records from ten data repositories used in our study. VR are records kept after data cleaning procedures. OBIS - Ocean Biogeographic Information System; GBIF - Global Biodiversity Information Facility; O'Hara *et al.* - reference in the main text; MNHN Paris - Muséum National d'Histoire Naturelle (France); ALA - Atlas of Living Australia; SIO - Scripps Institution of Oceanography (USA); Invemar - Instituto de Investigaciones Marinas y Costeras (Colombia); ICMYL - Instituto de Ciencias del Mar y Limnología (Mexico); Smithsonian NMNH - National Museum of Natural History (USA); NHM London - Natural History Museum (UK).

Repository	All Records	Valid Records (VR)	VR (%)
OBIS	200,714	80,228	57.81
GBIF	210,266	43,475	31.33
O'Hara <i>et al.</i>	15,537	9,573	6.90
MNHN Paris	10,949	2,272	1.64
ALA	31,730	1,790	1.29
SIO	1,483	668	0.48
Invemar	3,347	535	0.39
ICMYL	3,193	211	0.15
Smithsonian NMNH	4,000	17	0.01
NHM London	2,520	0	0

Supplementary Table 2 | Pearson product-moment correlation coefficients between latitudinal patterns in observed ophiuroids species richness, number of absent species, number of sampling events and estimates of inventory completeness between OBIS dataset and OBIS + additional datasets.

	<i>r</i>	<i>p</i> -value
N° of observed species	0.93	< 0.001
N° of species absent	0.99	< 0.001
N° of sampling events	0.98	< 0.001
Sousa-Baena' estimate	0.53	0.002
Sample coverage estimate	0.89	< 0.001
SACs estimate	0.94	< 0.001

Supplementary Table 3 | Number of records of all taxonomic group after each step of the data cleaning process (columns). Valid coords: marine records with both coordinates different from zero or within the sea boundary; Species level: records with species-level identification; Valid name: records with scientific name accepted by World Register of Marine Species; Valid date: records with sampling date available or, if unavailable, from a unique sampling event; S: total species richness. Data retrieved on February 22-23, 2017 for Ophiuroidea and April 11-17, 2017 for other taxa.

Taxon	Initial records	Valid coords	Species level	Excluding fossils	Excluding duplicates	Valid name	Valid date	S
Ophiuroidea	483739	441008	320890	319863	150420	148006	138769	1793
Bivalvia	974505	964471	823313	823101	435392	434302	422318	3789
Gastropoda	947776	910292	656780	651386	436285	428644	416105	11104
Copepoda	3698045	3682786	1988043	1987995	879273	875608	874254	3983
Porifera	378990	376956	151759	151516	101761	101611	100704	4902
Rhodophyta	422822	419328	321915	314032	184938	184856	181741	2555
Amphipoda	605049	599342	392659	392555	244344	243634	235322	4327
Scleractinia	588274	583904	472722	472318	135679	134512	122812	1382
Elasmobranchii	1082665	1078155	917292	917292	543644	543637	540319	962
Scombridae	890339	888193	862855	862855	411446	411444	410358	52
Fraction eliminated		1.79%	26.51%	0.28%	33.81%	0.19%	0.93%	

Supplementary Table 4 | Impact of removing records with double zero coordinates on estimates of species with spatial gaps. Records: number of records with double zero coordinates (latitude and longitude exactly equal to zero); Species level: number of records with species-level identification; S: number of species; S-Gap: number of species with spatial gaps; Total gaps: number of species with spatial gaps in the final filtered dataset; Impact: percentage of species with 0-0 location records removed that have a spatial gap in the final filtered dataset.

Taxon	Records	Species level	S	S-Gaps	Min depth	Max depth	Total gaps	Impact (%)
Ophiuroidea	2	2	1	0	11.75	14.65	814	0.00
Bivalvia	21	21	4	3	-	-	1459	0.21
Gastropoda	5	4	3	0	0	11.75	3397	0.00
Copepoda	1	1	1	1	-	-	1319	0.08
Porifera	0	0	0	0	-	-	1083	0.00
Rhodophyta	104	71	55	28	-	-	861	3.25
Amphipoda	15	15	14	10	250	250	1194	0.84
Scleractinia	3	3	3	1	8.84	8.84	808	0.12
Elasmobranchii	41	37	25	12	0	50.5	463	2.59
Scombridae	1	1	1	0	0	0	34	0.00

Supplementary References

- Chao, A. & Jost, L. (2012). Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology*, 93, 2533–2547.
- Herkt, K.M.B., Skidmore, A.K. & Fahr, J. (2017). Macroecological conclusions based on IUCN expert maps: A call for caution. *Glob. Ecol. Biogeogr.*, 26, 930–941.
- Hsieh, T.C., Ma, K.H. & Chao, A. (2016). iNEXT: iNterpolation and EXTrapolation for species diversity. R package version 2.0.12. Available at: <https://CRAN.R-project.org/package=iNEXT>
- Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., *et al.* (2018). vegan: community ecology package. R package version 2.5-3. Available at: <https://CRAN.R-project.org/package=vegan>
- Sousa-Baena, M.S., Garcia, L.C. & Peterson, A.T. (2014). Completeness of digital accessible knowledge of the plants of Brazil and priorities for survey and inventory. *Divers. Distrib.*, 20, 369–381.
- Stropp, J., Ladle, R.J., Malhado, A.C.M., Hortal, J., Gaffuri, J., H. Temperley, W., *et al.* (2016). Mapping ignorance: 300 years of collecting flowering plants in Africa. *Glob. Ecol. Biogeogr.*, 25, 1085–1096.
- Tittensor, D., Mora, C., Jetz, W., Lotze, H., Ricard, D., Vanden Berghe, E., *et al.* (2010). Global patterns and predictors of marine biodiversity across taxa. *Nature*, 466, 1098–1101.
- Vandepitte, L., Bosch, S., Tyberghein, L., Waumans, F., Vanhoorne, B., Hernandez, F., *et al.* (2015). Fishing for data and sorting the catch: assessing the data quality, completeness and fitness for use of data in marine biogeographic databases. *Database*, 2015, 1–14.
- Yang, W., Ma, K. & Kreft, H. (2013). Geographical sampling bias in a large distributional database and its effects on species richness-environment models. *J. Biogeogr.*, 40, 1415–1426.

Material suplementar

Capítulo 4

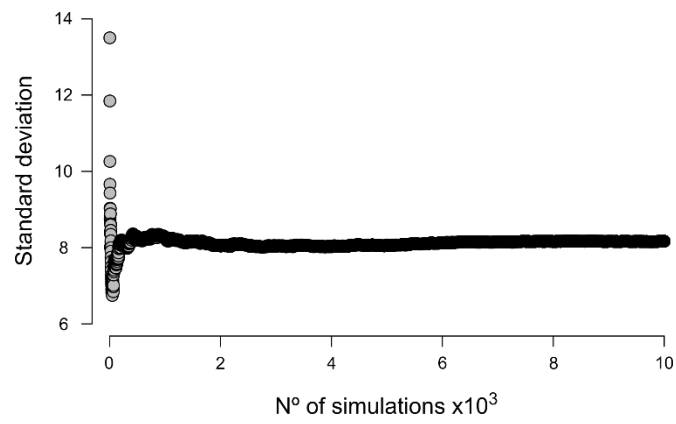


Figure S1. Model stability. Variation in standard deviation of the estimated diversity (i.e., sampled species richness) in a tropical latitude through 10,000 simulation replicates.

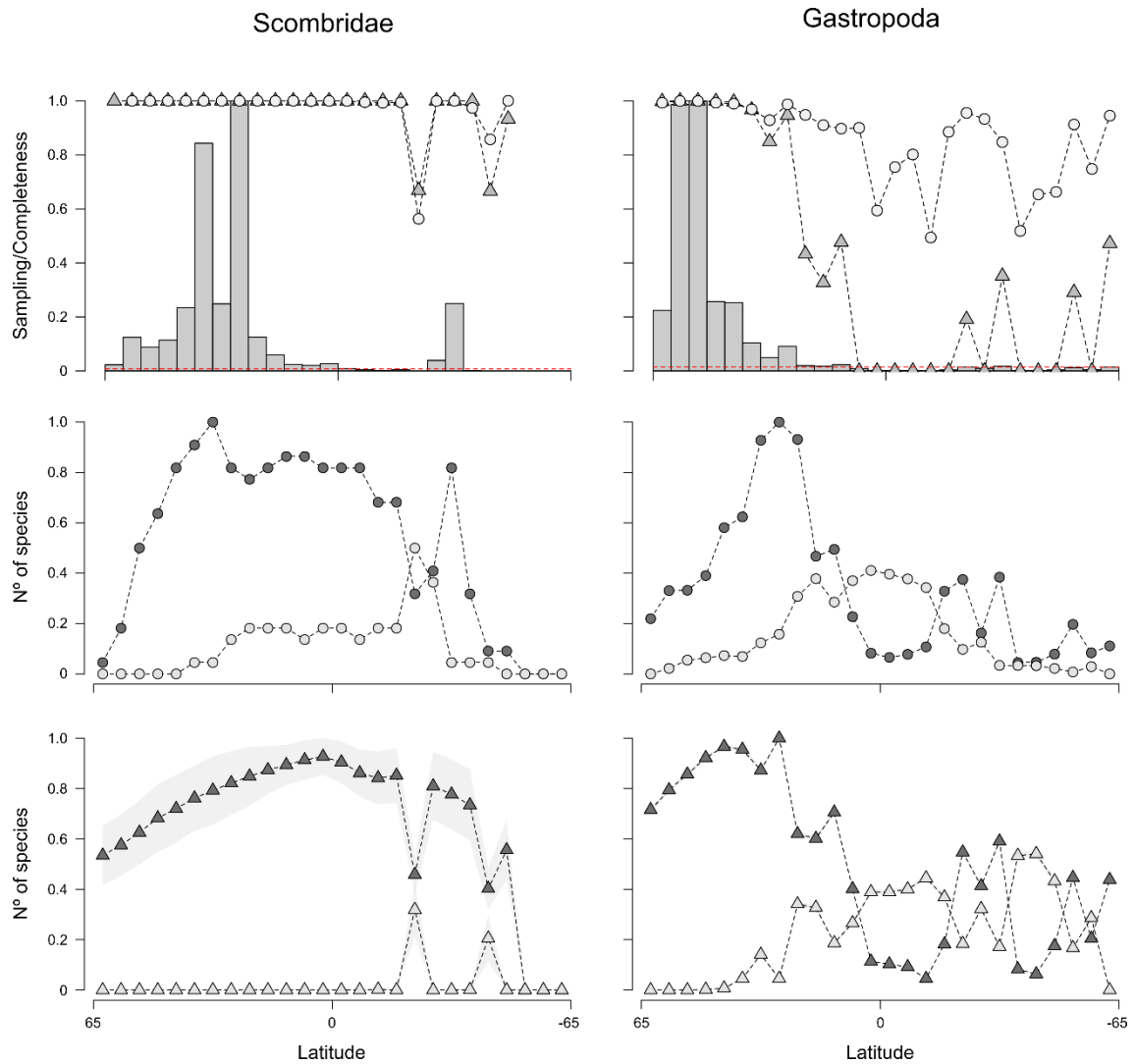


Figure S2. Inventory completeness and spatial gaps for tuna fishes and gastropods in the Atlantic Ocean. (Top) Inventory completeness estimated for the empirical data (light grey circles) and mid-domain model (grey triangles). Note in the gastropods plot how inventory completeness should be reduced if sampling effort were dilute through the latitudes. Histogram represent standardized sampling effort. For comparison purposes, we indicate the mark of 500 sampling events with the dotted red line. Plots below indicate the number of observed (dark grey) and missing species (light grey) in the empirical data (middle) and mid-domain model (bottom). Grey shadows represent standard deviation.

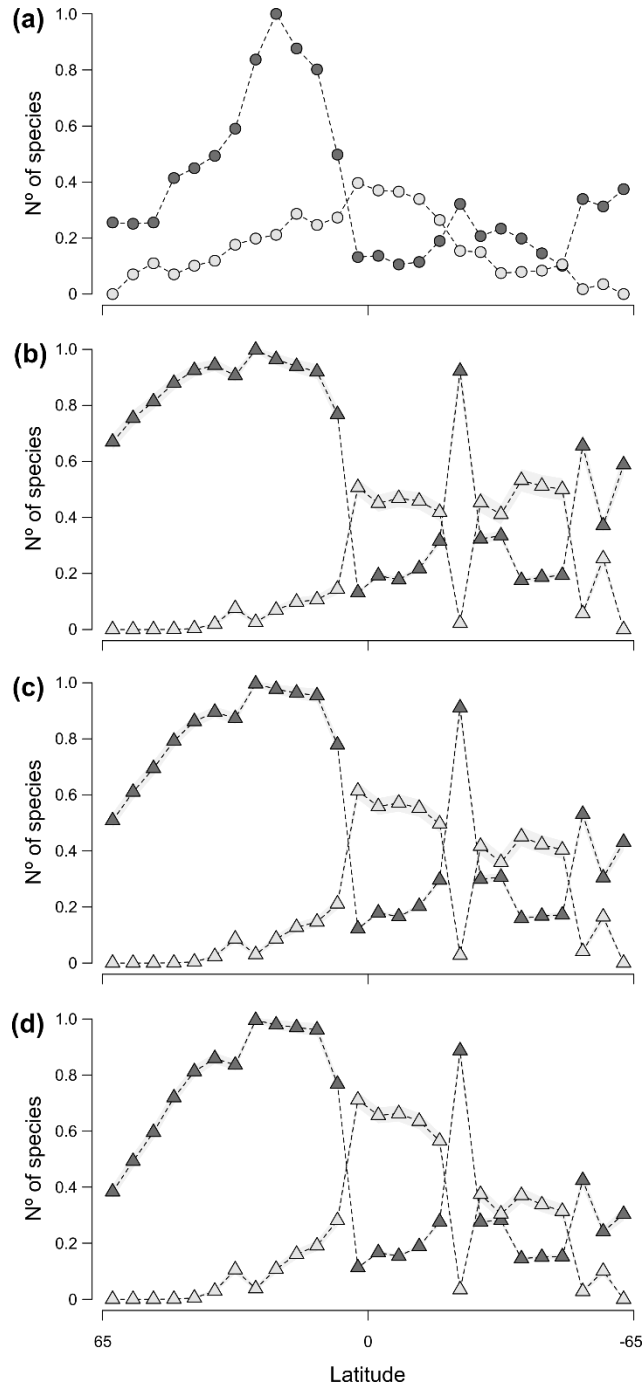


Figure S3. Tropical origin and species richness variation. Plots indicate the number of observed (dark grey) and missing species (light grey) in the empirical data (a) and in mid-domain models where species origination is draw randomly (b), or from a beta probability of $\alpha = \beta = 3$ (c) and $\alpha = \beta = 9$ (d). Note how the number of missing species is reduced in southern under-sampled latitudes as tropical origination rate increases. Grey shadows represent standard deviation.

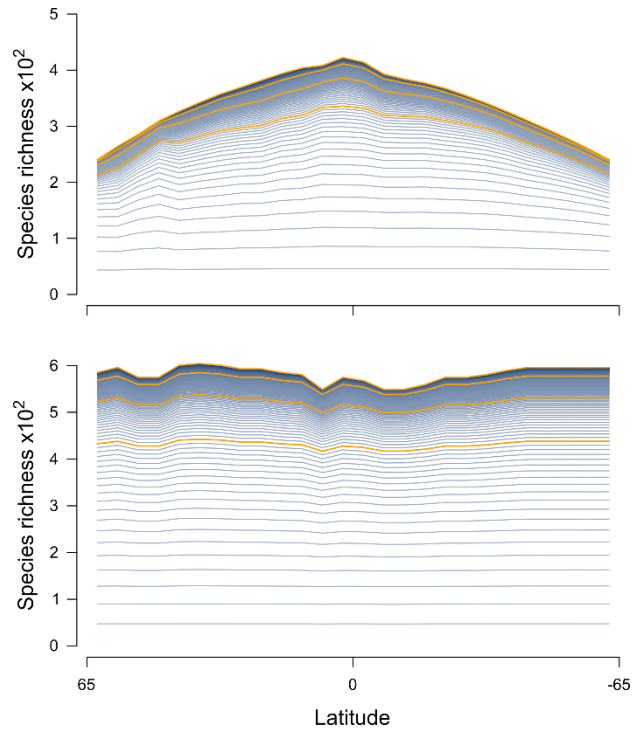


Figure S4. Species richness and sampling effort. Average number of species recorded by latitude in an increasing sampling effort simulation in the mid-domain (top) and area model (bottom). Orange lines indicate the observed species richness with 500, 1000, 2000 and 5000 sampling events at all latitudes.

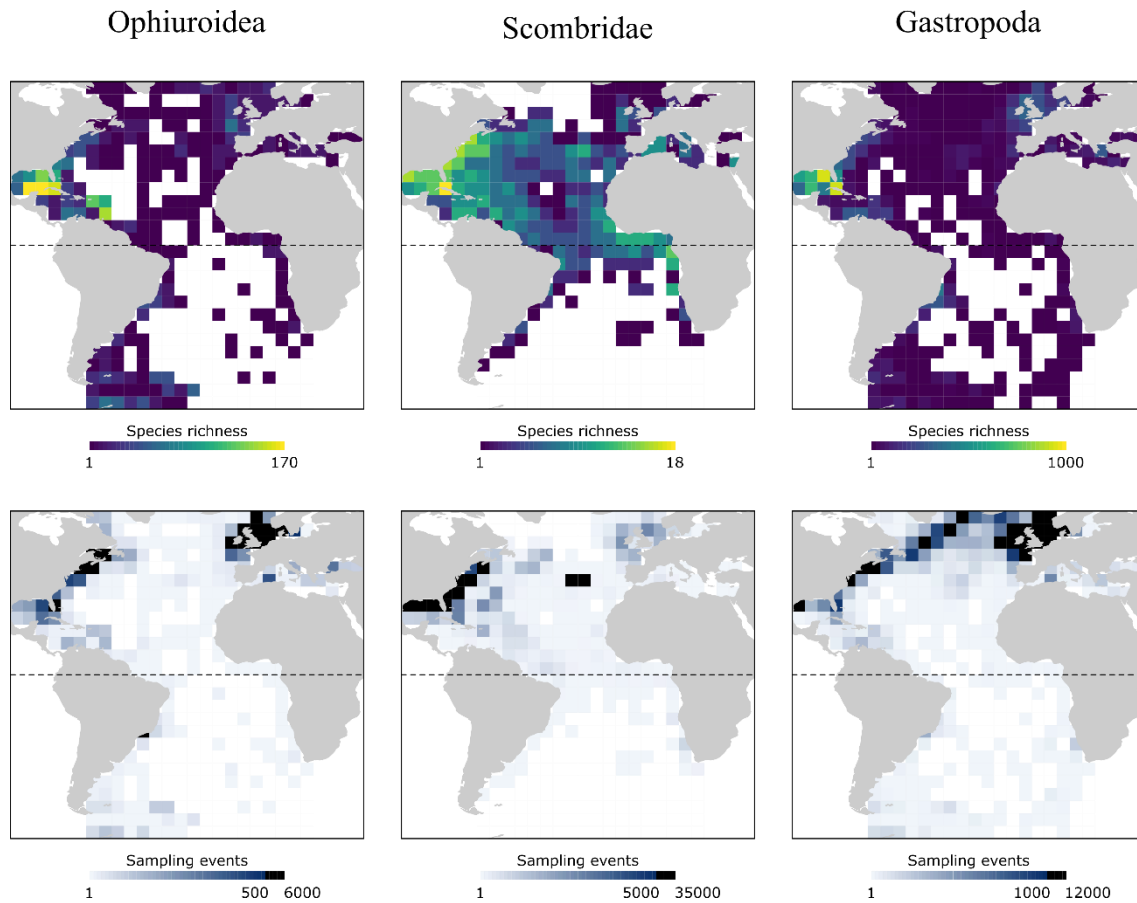


Figure S5. Empirical distribution of species richness and sampling effort in the Atlantic Ocean. Map showing the number of species (top) and sampling events (bottom) for ophiuroids, tuna fishes and gastropods. Note how the number of sampled cells is reduced in the southern hemisphere.

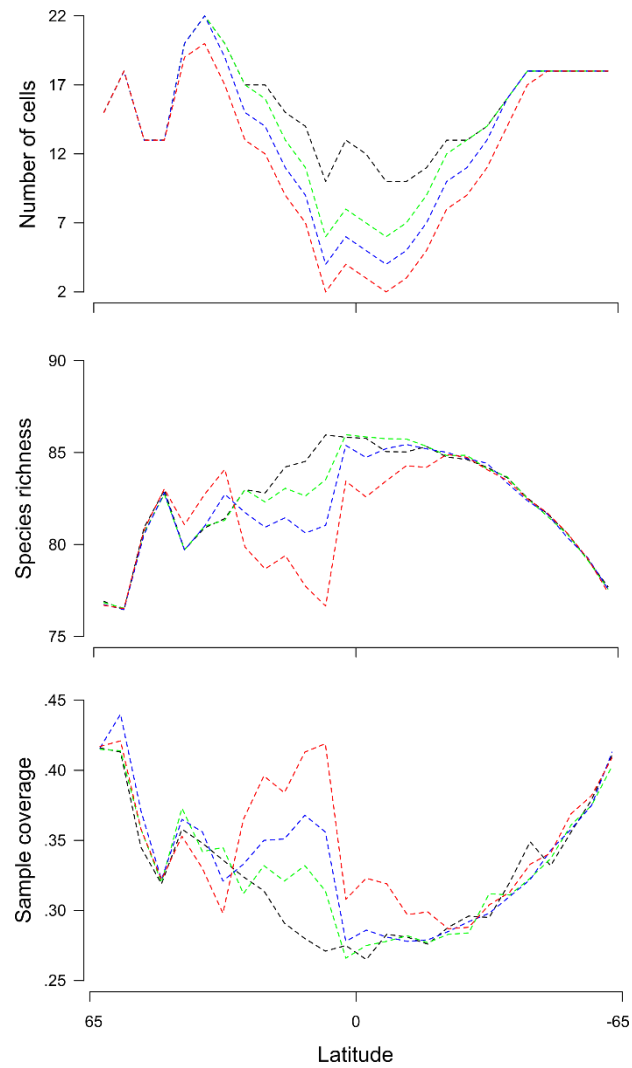


Figure S6. Effects of reduced sampling effort in tropical cells. Plots show the average number of observed species and the estimated sample coverage after 50 equal sampling events through the latitudes when all cells can be equally sampled (black dotted line) and when tropical cells available to sampling are reduced to a few options (green, blue and red dotted line).

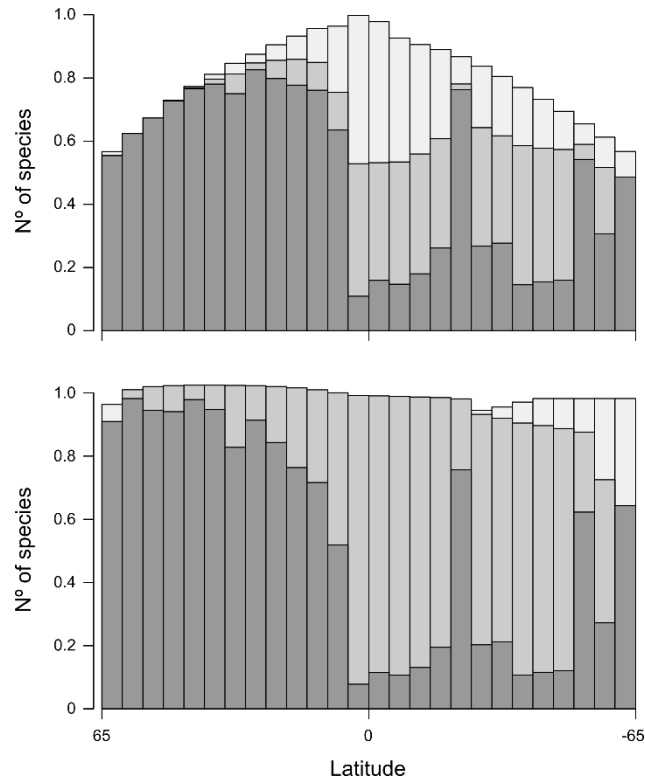


Figure S7. Filling the gaps. Histograms showing average values for the reference diversity per latitude (light grey) and the estimated number of species after sampling simulation (dark grey) for the mid-domain (top) and area model (bottom). Intermediate grey color represents the number of species that would be expected by summing the number of species recorded with those classified as “missing”. Note that even filling the gaps it is not possible to achieve the reference diversity. Data were standardized by the maximum species richness.

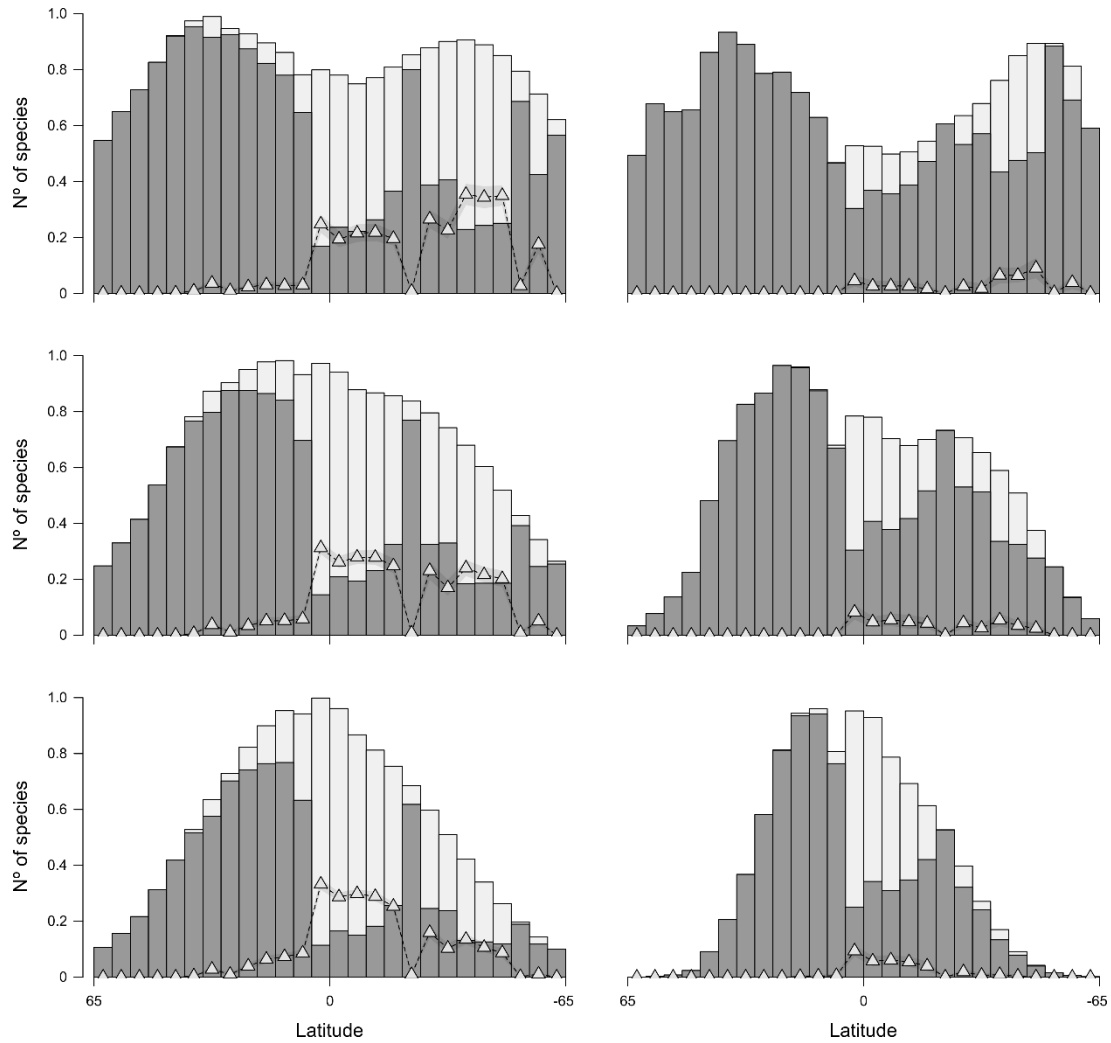


Figure S8. Reducing species geographical range. The beta distribution used to draw range sizes in our simulation results in ranges with $4 < 93 < 227$ cells (median between 1.5 interquartile range, calculated from 10,000 random observations). These histograms show the average number of species per latitude (light grey bars), the number of species recorded after the sampling simulation (dark grey bars) and the estimated number of missing species (i.e., spatial gaps; triangles) for the mid-domain model when range sizes are drawn from two extremely right skewed beta distributions ($\alpha = 1$, $\beta = 9$, $1 < 29 < 123$ cells, left panel; $\alpha = 1$, $\beta = 100$, $1 < 3 < 12$ cells, right panel). Plots show the resulting latitudinal diversity when species origination cell is draw randomly (top), or from a given latitude with beta probability of $\alpha = \beta = 3$ (middle) and $\alpha = \beta = 9$ (bottom). Data were standardized by the maximum species richness. Note how diversity pattern becomes bimodal as mean range size decreases. Reducing range size also reduce the number of species per latitude, which increases sampling efficiency. However, the frequency of spatial gaps is also highly reduced in such situation. Overall, assuming a cohesive range, high frequency of tropical gaps only occurs in a scenario with high tropical richness and not so small geographic ranges. Applying these range size frequency distributions to the area model resulted in a diversity pattern similar to that in the top-right plot, but with spatial gaps unrealistically four-fold higher than reference diversity in lower latitudes (results not shown).