

## Area-based conservation planning in Japan: The importance of OECMs in the post-2020 Global Biodiversity Framework

Takayuki Shiono <sup>a,1,\*</sup>, Yasuhiro Kubota <sup>a,b,2,\*</sup>, Buntarou Kusumoto <sup>c,3</sup>

<sup>a</sup> Faculty of Science, University of the Ryukyus, Okinawa, Japan

<sup>b</sup> Think Nature Inc., Okinawa, Japan

<sup>c</sup> Faculty of Agriculture, Kyushu University, Japan



### ARTICLE INFO

**Keywords:**

Connectivity  
Land sharing/sparing  
OECM  
Spatial prioritization  
Satoyama  
Zonation

### ABSTRACT

To reframe the imperfect review processes of nation-scale actions on area-based conservation through protected area (PA) networks, we first created novel infrastructure to visualize nation-level biodiversity information in Japan. We then assessed the performance of the existing PA network relative to land exploitation pressure and evaluated conservation effectiveness of PA expansion for the post-2020 Global Biodiversity Framework. The Zonation software was used to spatially prioritize conservation areas to minimize biodiversity loss and the extinction risk for 8077 Japanese vascular plant and vertebrate species under constraints of the existing PA network and land use. The spatial pattern of the identified priority areas, which were considered candidate areas for expansion of the current PA network, was influenced by land-use types according to the mask layers of non-PAs, and low-, middle-, and high-ranked PAs. The current PA network reduced the aggregate extinction risk of multiple species by 36.6%. Indeed, the proportion of built-up areas in the existing PAs was generally smaller than that in the areas surrounding PAs. Notably, high-ranked PAs effectively restricted the built-up pressure (0.04% every 10 years), whereas low-ranked PAs in national park and wild-life protection areas did not (1.8% every 10 years). Conservation effects were predicted to substantially improve by expansion of high-ranked (legally strict) PAs into remote non-PAs without population/socioeconomic activities, or expansion of medium-ranked PAs into satoyama which have traditionally been used for agriculture and forestry, and urban areas. A 30% land conservation target was predicted to decrease extinction risk by 74.1% when PA expansion was implemented across remote areas, satoyama, and urban areas; moreover, PA connectivity almost doubled compared with the existing PA network. In contrast, a conventional scenario showed that placing national parks in state-owned and non-populated areas would reduce extinction risk by only 4.0%. Conservation prioritization analyses demonstrated the effectiveness of using a comprehensive conservation approach that reconciles land-sparing protection and land-sharing conservation in other effective area-based conservation measures (OECMs) in satoyama and urban green spaces. Our results revealed that complementary inclusion of various PA interventions related to their governance and land-use planning plays a critical role in effectively preventing biodiversity loss and makes it more feasible to achieve ambitious conservation targets.

\* Corresponding authors at: Faculty of Science, University of the Ryukyus, Okinawa, Japan.

E-mail addresses: [biodiver.shiono@gmail.com](mailto:biodiver.shiono@gmail.com) (T. Shiono), [kubota.yasuhiro@gmail.com](mailto:kubota.yasuhiro@gmail.com) (Y. Kubota).

<sup>1</sup> ORCID ID (0000-0001-7253-7956)

<sup>2</sup> ORCID ID (0000-0002-5723-4962)

<sup>3</sup> ORCID ID (0000-0002-5091-3575)

<https://doi.org/10.1016/j.gecco.2021.e01783>

Received 5 July 2021; Received in revised form 26 August 2021; Accepted 26 August 2021

Available online 27 August 2021

2351-9894/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**SPECIAL FEATURE**

Approaches for general rules of biodiversity patterns in space and time

# Global distribution of coral diversity: Biodiversity knowledge gradients related to spatial resolution

Buntarou Kusumoto<sup>1,2,3</sup> | Mark J. Costello<sup>4</sup> | Yasuhiro Kubota<sup>2,5</sup> |  
 Takayuki Shiono<sup>2</sup> | Chi-Lin Wei<sup>6</sup> | Moriaki Yasuhara<sup>7</sup> | Anne Chao<sup>8</sup>

<sup>1</sup>Royal Botanic Gardens, Kew, Richmond, UK

<sup>2</sup>Faculty of Science, University of the Ryukyus, Okinawa, Japan

<sup>3</sup>Okinawa Prefecture Environment Science Center, Okinawa, Japan

<sup>4</sup>School of Environment, The University of Auckland, Auckland, New Zealand

<sup>5</sup>Marine and Terrestrial Field Ecology, Tropical Biosphere Research Center, University of the Ryukyus, Okinawa, Japan

<sup>6</sup>Institute of Oceanography, National Taiwan University, Taipei, Taiwan

<sup>7</sup>School of Biological Sciences and Swire Institute of Marine Science, The University of Hong Kong, Hong Kong, China

<sup>8</sup>Institute of Statistics, National Tsing Hua University, Hsin-Chu, Taiwan

## Correspondence

Buntarou Kusumoto, Faculty of Science, University of the Ryukyus, Okinawa, Japan.

Email: kusumoto.buntarou@gmail.com

## Funding information

JSPS KAKENHI, Grant/Award Number: 16H02755; University of the Ryukyus; Program for Advancing Strategic International Networks to Accelerate the Circulation of Talented Researchers

## Abstract

Biodiversity knowledge shortfalls, especially incomplete information on species distributions, can lead to false conclusions about global biodiversity patterns. Diversity estimation theory statistically uses species occurrence records and sampling completeness (coverage) to predict diversity in terms of species richness, dominance and evenness. We estimated Scleractinia coral species diversity at different spatial resolutions, based on 109,296 occurrences and range data for 697 species, using an incidence-based Hill's numbers approach through a rarefaction and extrapolation technique. We found that spatial patterns of diversity estimates were dependent on a geographic scale. The latitudinal and longitudinal diversity gradients, particularly at finer spatial scales, differed from species range-based coral biodiversity hotspots of previous studies. The western Indian Ocean was predicted to have the most coral species, with greater diversities than in the Indo-Pacific Coral Triangle. We concluded that the identification of marine biodiversity hotspots is sensitive to species commission errors (from range maps) and biased sampling coverage. Moreover, estimates of the geographic distribution of species richness informed us of a set of priority areas (the northeastern coast of Australia, central Coral Triangle and coast of Madagascar) for future sampling of unknown coral species occurrence. Our findings of biogeographical survey priorities contribute to filling biodiversity shortfalls for tropical coral reefs through sampling completeness, and consequently for development of conservation planning.

## KEY WORDS

coral reefs, Hill's number, species incidence, sampling bias, marine biodiversity hotspots

## 1 | INTRODUCTION

Fine-scale biodiversity mapping is fundamental to macroecological studies and spatial conservation prioritization (Gaston, 2000; Jetz, McPherson, & Guralnick, 2012;

Moilanen, Wilson, & Possingham, 2009; Tittensor et al., 2010). In recent decades, this approach has led to the creation of databases referencing the occurrence of millions of species (or specimen records), which are potentially significant for biodiversity assessment in space and time

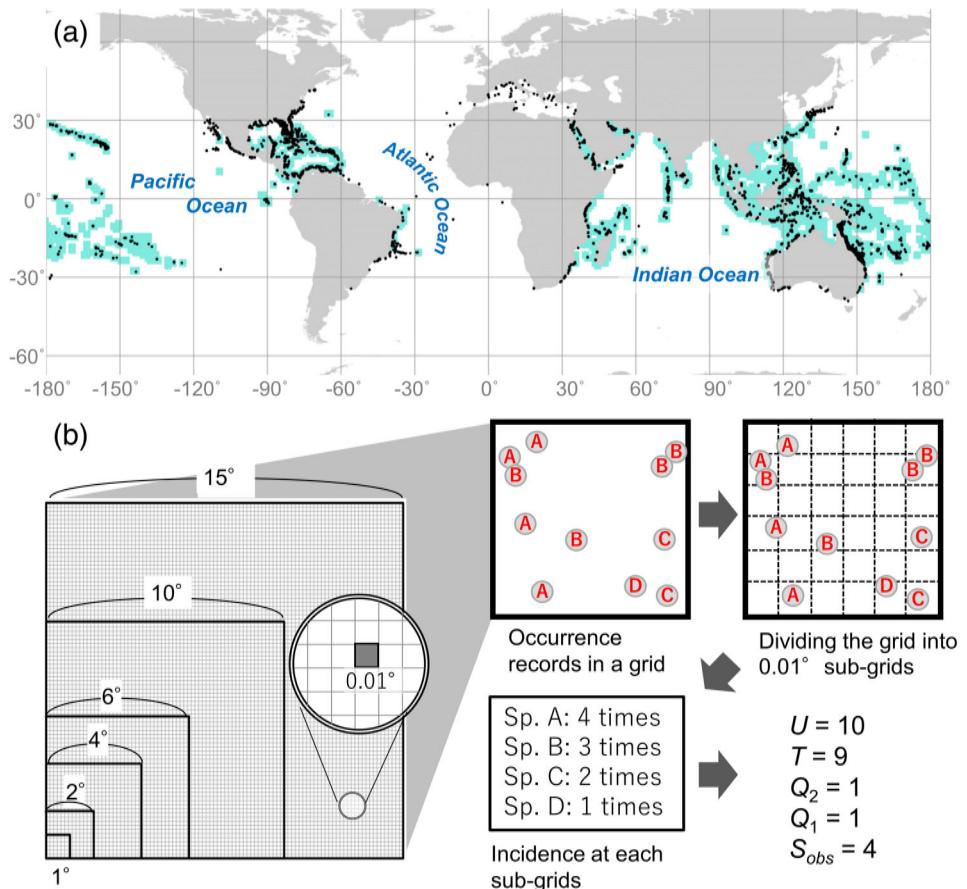
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Ecological Research* published by John Wiley & Sons Australia, Ltd on behalf of The Ecological Society of Japan

[www.iobis.org/](http://www.iobis.org/)). To compile a data set of species occurrence, we excluded occurrence records not identified to species level and/or without a geographical coordinate. We also removed duplicated information (i.e., identical species names, longitudes and latitudes, where the averaged positioning resolution was approximately  $0.2^{\circ}$ ) within and between the two data sources. Species names were standardized following the World Register of Marine Species (WoRMS) (Horton et al., 2019); subspecies and variety were merged into their respective species. To focus on tropical coral reefs, we filtered out particular species (e.g., deep-sea or cold-water corals) not previously recorded from coral reef areas (Figure 1; ReefBase; <http://www.reefbase.org/main.aspx>). We finally obtained 109,296 geo-referenced occurrence points for 697 species (Figure 1a).

## 2.2 | Defining species incidence

We defined the frequency of species incidence at six spatial resolutions: 1, 2, 4, 6, 10 and  $15^{\circ}$  grid cells (Figure 1b). We divided each grid cell into  $0.01^{\circ}$  of sub-gridded cells, counted the number of sub-gridded cells that contained occurrence records for individual species, and then created the dataset of species incidence at the grid cell level. To obtain reliable estimates of species diversity, we excluded grid cells with few occurrence records from analysis (Figure S1), that is, if the observed number of species was less than six, the number of sub-gridded cells with at least one incidence was less than six, or the total number of species incidences was equal to the number of unique species (species that are each detected in only one sub-grid cell).



**FIGURE 1** Global distribution of occurrence records of tropical Scleractinia stony corals and the evaluation procedure of species incidence to estimate species diversity. (a) Occurrence records of coral species derived from the Global Biodiversity Information Facility and Ocean Biogeographic Information System (109,367 points). Light-blue areas represent the coral reef distribution provided by ReefBase (<http://www.reefbase.org/main.aspx>). (b) Compilation of species incidence data: dividing the globe into grids of different size of (1, 2, 4, 6, 10 and  $15^{\circ}$ ); subdividing each grid into  $0.01 \times 0.01$  sub-gridded cells; counting the number of sub-gridded cells containing occurrence records in each grid for each species; creating species incidence distribution in each grid and calculating relevant values ( $U$  = sum of species incidences,  $T$  = number of sub-grids where at least one incidence was found,  $Q_2$  = number of duplicates,  $Q_1$  = number of uniques,  $S_{obs}$  = observed number of species)

ARTICLE



<https://doi.org/10.1038/s41467-020-15407-5>

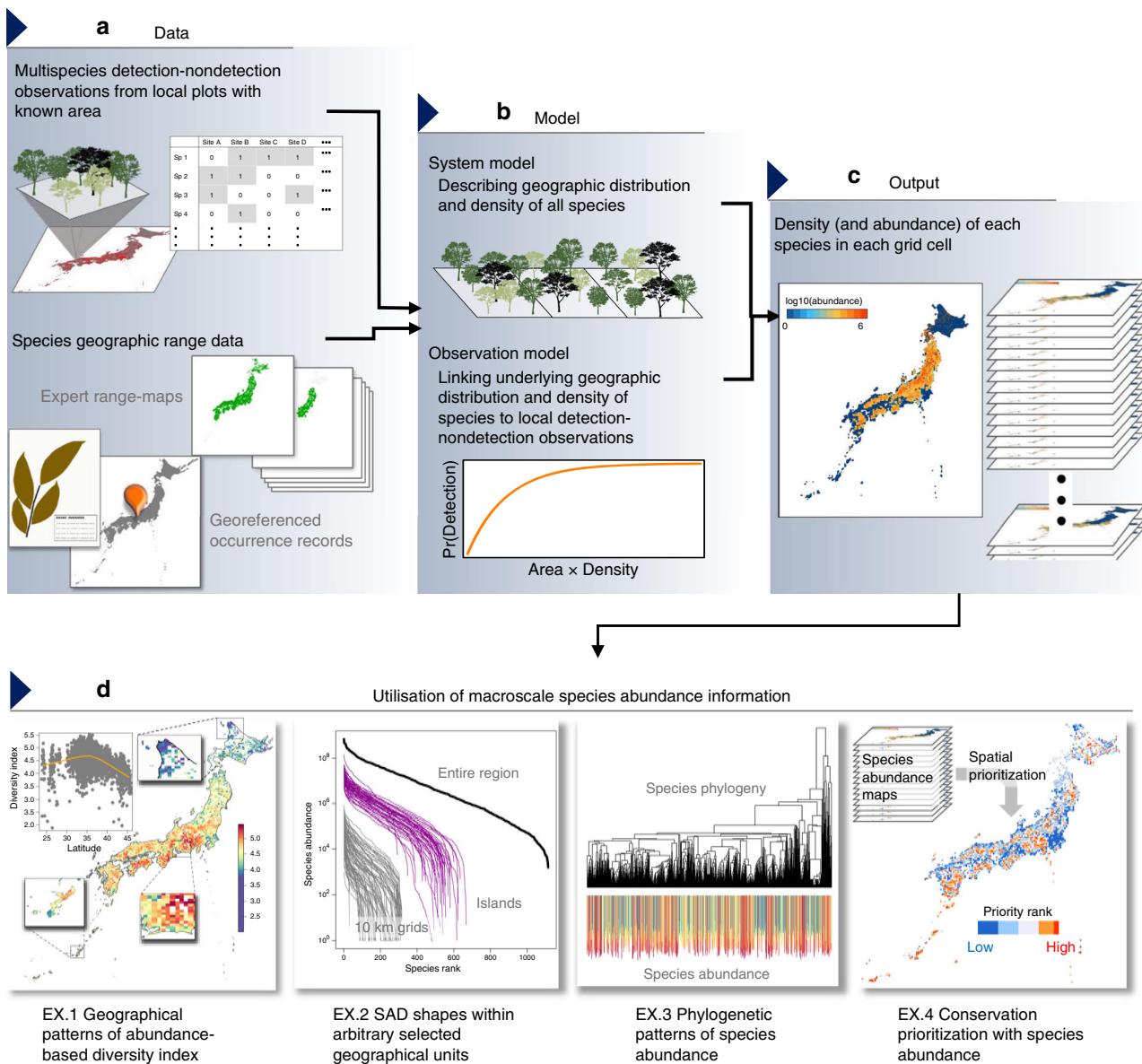
OPEN

# Integrating multiple sources of ecological data to unveil macroscale species abundance

Keiichi Fukaya  <sup>1,2</sup>✉, Buntarou Kusumoto  <sup>3</sup>, Takayuki Shiono  <sup>3</sup>, Junichi Fujinuma  <sup>3</sup> & Yasuhiro Kubota  <sup>3</sup>

The pattern of species abundance, represented by the number of individuals per species within an ecological community, is one of the fundamental characteristics of biodiversity. However, despite their obvious significance in ecology and biogeography, there is still no clear understanding of these patterns at large spatial scales. Here, we develop a hierarchical modelling approach to estimate macroscale patterns of species abundance. Using this approach, estimates of absolute abundance of 1248 woody plant species at a 10-km-grid-square resolution over East Asian islands across subtropical to temperate biomes are obtained. We provide two examples of the basic and applied use of the estimated species abundance for (1) inference of macroevolutionary processes underpinning regional biodiversity patterns and (2) quantitative community-wide assessment of a national red list. These results highlight the potential of the elucidation of macroscale species abundance that has thus far been an inaccessible but critical property of biodiversity.

<sup>1</sup>National Institute for Environmental Studies, 16-2 Onogawa, Tsukuba, Ibaraki 305-8506, Japan. <sup>2</sup>The Institute of Statistical Mathematics, 10-3 Midoricho, Tachikawa, Tokyo 190-8562, Japan. <sup>3</sup>Faculty of Science, University of the Ryukyus, 1 Senbaru, Nishihara, Okinawa 903-0213, Japan. ✉email: [fukaya\\_keiichi@nies.go.jp](mailto:fukaya_keiichi@nies.go.jp)



**Fig. 1 A framework for estimation of macroscale species abundance.** Spatially replicated detection-nondetection observations and various information on species geographic distribution **a** are integrated in a hierarchical model that links binary observations to underlying species abundance **b**. A model fitting yields estimates of individual density of each species in each geographic grid cell, which can then be used to derive estimates of species abundance with the area of suitable habitat **c**. The results can be used for diverse purposes relevant to, e.g. community ecology, macroecology, biogeography, and applied fields of ecology **d**.

of each species are provided in Supplementary Note 4. In the following subsections, we describe two post hoc analyses that highlight the utility of macroscale species abundance estimates.

**Inferring macroevolutionary processes of biodiversity.** The UNTB predicts that the statistical form of the SAD in the metacommunity is intimately linked to the mode of speciation<sup>5,11,12,18–20</sup>. This theoretical foundation enables utilizing the UNTB to infer the role of macroevolution in shaping ecological patterns based on statistical analyses of SADs. Nevertheless, the fact that species abundance data can be obtained only from local communities has been a critical limitation to the practical inference of macroevolutionary processes<sup>5,11,12,18</sup>. Preparing estimates of macroscale species abundance may prove to be a useful and probably only solution to this problem, given that

obtaining data on species abundance over a huge spatial extent is obviously unrealistic.

We obtained metacommunity SADs for four ecoregions that belong to different biogeographic groups (i.e. the central continental arc, northern continental arc, southern continental arc, and oceanic islands; Fig. 4) by aggregating abundance estimates over grid cells within each region. For each ecoregion, three variants of the UNTB, the point mutation speciation model<sup>5,21</sup>, random fission speciation model<sup>12</sup>, and protracted speciation model<sup>11</sup>, were fitted to make an inference about the metacommunity SAD. Additional details of this analysis are described in the “Methods” section (subsection, Inference of macroevolutionary processes in metacommunities).

The SADs of metacommunities in the four ecoregions followed a left-skewed, lognormal-like distribution, whose short left tail indicates that the number of very rare species was negligible (Fig. 4). Among the three variants of the UNTB, this pattern of

## Article

# Islands are key for protecting the world's plant endemism

<https://doi.org/10.1038/s41586-024-08036-1>

Received: 28 February 2024

Accepted: 11 September 2024

Published online: 16 October 2024



Julian Schrader<sup>1</sup>✉, Patrick Weigelt<sup>2,3,4</sup>, Lirong Cai<sup>2</sup>, Mark Westoby<sup>1</sup>, José María Fernández-Palacios<sup>5</sup>, Francisco J. Cabezas<sup>6</sup>, Gregory M. Plunkett<sup>7</sup>, Tom A. Ranker<sup>8</sup>, Kostas A. Triantis<sup>9</sup>, Panayiotis Trigas<sup>10</sup>, Yasuhiro Kubota<sup>11</sup> & Holger Kreft<sup>2,3,4</sup>

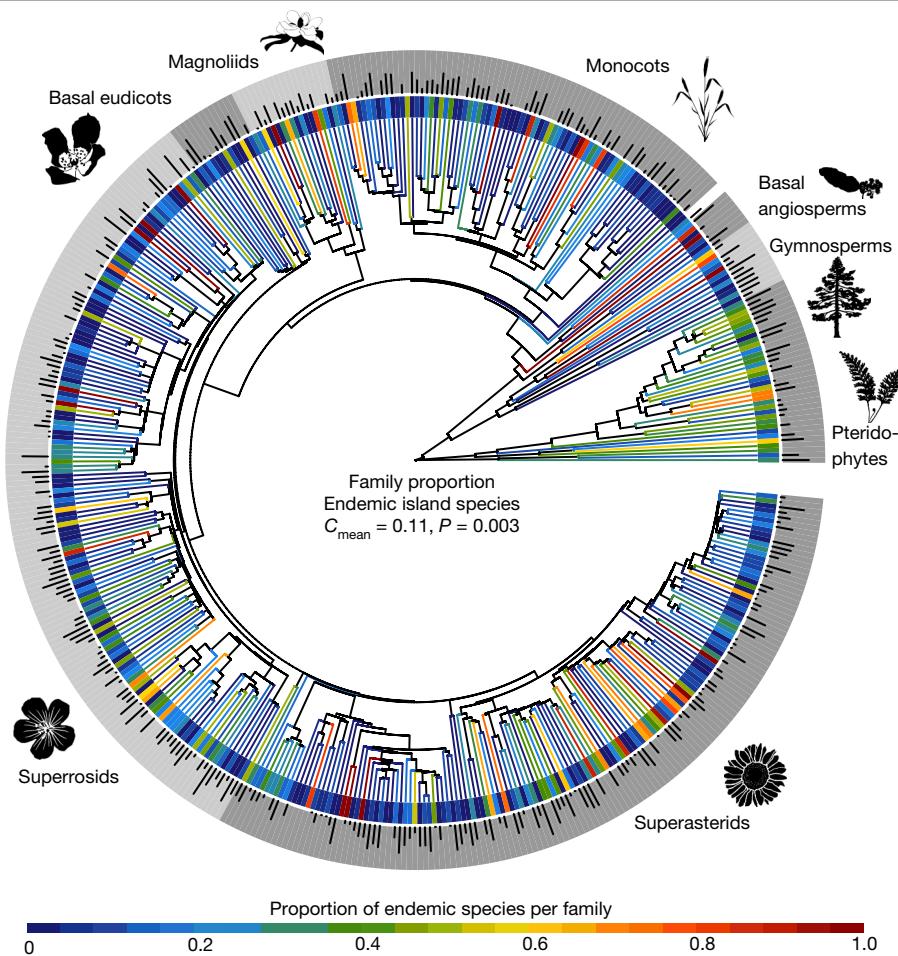
Islands are renowned as evolutionary laboratories and support many species that are not found elsewhere<sup>1,2</sup>. Islands are also of great conservation concern, with many of their endemic species currently threatened or extinct<sup>3</sup>. Here we present a standardized checklist of all known vascular plants that occur on islands and document their geographical and phylogenetic distribution and conservation risk. Our analyses of 304,103 plant species reveal that 94,052 species (31%) are native to islands, which constitute 5.3% of the global landmass<sup>4</sup>. Of these, 63,280 are island endemic species, which represent 21% of global plant diversity. Three-quarters of these are restricted to large or isolated islands. Compared with the world flora, island endemics are non-randomly distributed within the tree of life, with a total of 1,005 billion years of unique phylogenetic history with 17 families and 1,702 genera being entirely endemic to islands. Of all vascular plants assigned International Union for Conservation of Nature conservation categories<sup>5</sup>, 22% are island endemics. Among these endemic species, 51% are threatened, and 55% of all documented global extinctions have occurred on islands. We find that of all single-island endemic species, only 6% occur on islands meeting the United Nations 30×30 conservation target. Urgent measures including habitat restoration, invasive species removal and ex situ programmes are needed to protect the world's island flora. Our checklist quantifies the uniqueness of island life, provides a basis for future studies of island floras, and highlights the urgent need to take actions for conserving them.

Islands have served as natural laboratories for studying assembly processes, speciation and ecological adaptation<sup>2</sup>. The evolution of life in isolation has led to unique species assemblages and remarkable examples of diversification and adaptive radiations, including morphological oddities and island syndromes such as the tendency of herbaceous plants to become woody or the loss of defence mechanisms<sup>6,7</sup>. In particular, islands support many species that are not found anywhere else<sup>1,8</sup> and are hotspots of phylogenetic endemism<sup>9</sup>. Extrapolations have suggested that as many as 70,000 plant species worldwide could be endemic to islands<sup>1</sup>. These include species such as the New Caledonian endemic *Amborella trichopoda*, a lineage that forms the extant sister to all other angiosperms<sup>10</sup> or the palm genus *Howea*, which consists of two species that evolved sympatrically on Lord Howe Island, Australia<sup>11</sup>. Island endemics also feature classic examples of species radiations, such as the 63 extant members of the *Aeonium* alliance in Macaronesia<sup>12</sup> or the 126 species of lobeliads that are endemic to Hawai'i<sup>13</sup>, both of which radiated from single ancestors. Islands also harbour some of the rarest and most threatened taxa worldwide. Examples include *Brighamia insignis*, a species that is endemic to the Kaua'i and Ni'ihau

islands in Hawai'i, was last seen in the wild in 2012<sup>14</sup> and is now considered extinct<sup>5</sup>, or the rare ebony tree *Diospyros egrettarum*, whose only known viable population is from the 25-ha Île aux Aigrettes near Mauritius<sup>15</sup>. Despite their value for evolutionary and ecological research and conservation, and their many uses for humans, great uncertainty remains around estimates of the total number of plants native and endemic to islands and their global distribution. Thus, the current lack of a working list of all plant species that are known to be native or endemic to islands and their distribution status, which should form the basis of such a global assessment, presents a major knowledge gap. This is concerning because islands are at the forefront of biodiversity loss. Many of their unique species are threatened by extinction, which also endangers crucial services that are provided by island ecosystems, including cultural and spiritual values, and increases the vulnerability of Indigenous people and local communities<sup>3,16,17</sup>.

Here we provide a comprehensive assessment of all vascular plant species native and endemic to marine islands worldwide and highlight their conservation risk. Based on 5,243 taxonomically standardized checklists and floras for 1,967 island and 1,010 mainland regions covering the

<sup>1</sup>School of Natural Sciences, Macquarie University, Sydney, New South Wales, Australia. <sup>2</sup>Department of Biodiversity, Macroecology and Biogeography, University of Göttingen, Göttingen, Germany. <sup>3</sup>Centre of Biodiversity and Sustainable Land Use, University of Göttingen, Göttingen, Germany. <sup>4</sup>Campus Institute Data Science, University of Göttingen, Göttingen, Germany. <sup>5</sup>Island Ecology and Biogeography Research Group, Instituto Universitario de Enfermedades Tropicales y Salud Pública de Canarias (IUETSPC), Universidad de La Laguna, Tenerife, Spain. <sup>6</sup>Department of Biodiversity, Ecology and Evolution, Faculty of Biological Sciences, Complutense University of Madrid, Madrid, Spain. <sup>7</sup>Cullman Program for Molecular Systematics, New York Botanical Garden, New York, NY, USA. <sup>8</sup>School of Life Sciences, University of Hawai'i at Mānoa, Honolulu, HI, USA. <sup>9</sup>Department of Ecology and Taxonomy, Faculty of Biology, National and Kapodistrian University, Athens, Greece. <sup>10</sup>School of Plant Sciences, Agricultural University of Athens, Athens, Greece. <sup>11</sup>Faculty of Science, University of the Ryukyus, Nishihara, Japan. <sup>✉</sup>e-mail: jschrader@posteo.de



**Fig. 2 | Proportion of endemic island plants at family level.** Colours indicate proportions of endemic species per family. Tip and inner ring colours indicate coverage of endemic island plants at family level and the outer ring delimits major clades. Bar heights in the outer ring are proportional to  $\log_{10}$ -transformed richness of endemic species per family. Proportions of island endemics are significantly different between families. Phylogenetic signal down to family level

is expressed as Abouheif's  $C_{\text{mean}}$ , a measure of phylogenetic autocorrelation that is calculated by summing the squared differences between values of adjacent tips in the phylogeny; the significant  $P$  value indicates that families closer to each other in the phylogeny tend to be more similar in their proportion of island endemics. Plant silhouettes were created with PhyloPic.org.

contribution of islands to the plant tree of life. However, island endemics are non-randomly distributed across families compared to the world flora. Families that are phylogenetically closer tended to have more similar percentages of island endemics ( $C_{\text{mean}} = 0.11, P = 0.003$ ; Fig. 2). This means that the tendency to colonize islands and to radiate there is to some degree phylogenetically conserved both at family level and above.

Endemism on islands is hypothesized to increase with island area and isolation through their combined effect on speciation rates<sup>36–38</sup>. Larger islands have lower extinction rates due to larger population sizes and lower vulnerabilities to catastrophic events, increasing the lineages' survival long enough to become differentiated from the mainland populations, thus enhancing the chances of lineages to evolve into a new species<sup>37</sup>. Larger islands also provide a wider variety of habitats and feature more barriers to gene flow, which further increases speciation rates<sup>39</sup>. Smaller islands, which support smaller population sizes, have higher natural extinction rates<sup>40</sup>, making in situ speciation less likely. Isolation limits gene flow via inter-island dispersal and colonization from the mainland, leading to higher speciation rates<sup>41</sup>, and increases endemism by having fewer arrivals of new competitors compared with islands near mainland<sup>35</sup>. Although other factors such as island age, ontogeny and past climate are also hypothesized to be important influences on island endemism<sup>21,42</sup>, area and isolation show the highest predictive power<sup>21,43</sup>.

Endemism on islands and archipelagos globally is best explained by the interaction of island area and isolation (Fig. 3a), especially when accounting for island type and archipelago configuration ( $R^2 = 0.74$ ; see Methods for model details). Considerable variation in the numbers of endemics is explained by island area alone (continental:  $R^2 = 0.55$ ; complex and fragment:  $R^2 = 0.54$ ; oceanic:  $R^2 = 0.19$ ; Extended Data Fig. 4a,c), since larger islands carry more species in total, including endemics<sup>43</sup>. At island and archipelago level, isolation alone had only a weak influence ( $R^2$  values  $\leq 0.05$ ; Fig. 3 and Extended Data Fig. 4b,d). For oceanic archipelagos, this is probably because the most isolated ones, such as Hawai'i, Galápagos or Fiji, share many endemic species with other islands of the same archipelago owing to intra-archipelago dispersal, limiting the total numbers and proportions of single-island endemics. At the archipelago level (Fig. 3b), many of the most isolated archipelagos comprise small atolls, such as Tuamotu, Gilbert and Line Islands in the Pacific, or the Maldives and Chagos Archipelago in the Indian Ocean, which support a widespread and largely non-endemic strand flora<sup>44</sup>. This diminished the overall effect of isolation on endemism.

### Threat status of island plants

Of 36 recognized global biodiversity hotspots<sup>45</sup>, 9 are located exclusively on islands and 3 others include a substantial share of islands.

## RESEARCH ARTICLE

# Multiple drivers of the COVID-19 spread: The roles of climate, international mobility, and region-specific conditions

Yasuhiro Kubota  <sup>1,2\*</sup>, Takayuki Shiono  <sup>1</sup>, Buntarou Kusumoto  <sup>3</sup>, Junichi Fujinuma  <sup>1</sup>

**1** Faculty of Science, University of the Ryukyus, Okinawa, Japan, **2** Think Nature Inc., Okinawa, Japan, **3** University Museum, University of the Ryukyus, Okinawa, Japan

\* [Kubota.yasuhiro@gmail.com](mailto:Kubota.yasuhiro@gmail.com)



## OPEN ACCESS

**Citation:** Kubota Y, Shiono T, Kusumoto B, Fujinuma J (2020) Multiple drivers of the COVID-19 spread: The roles of climate, international mobility, and region-specific conditions. PLoS ONE 15(9): e0239385. <https://doi.org/10.1371/journal.pone.0239385>

**Editor:** Yury E. Khudyakov, Centers for Disease Control and Prevention, UNITED STATES

**Received:** May 18, 2020

**Accepted:** September 5, 2020

**Published:** September 23, 2020

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0239385>

**Copyright:** © 2020 Kubota et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

## Abstract

Following its initial appearance in December 2019, coronavirus disease 2019 (COVID-19) quickly spread around the globe. Here, we evaluated the role of climate (temperature and precipitation), region-specific COVID-19 susceptibility (BCG vaccination factors, malaria incidence, and percentage of the population aged over 65 years), and human mobility (relative amounts of international visitors) in shaping the geographical patterns of COVID-19 case numbers across 1,020 countries/regions, and examined the sequential shift that occurred from December 2019 to June 30, 2020 in multiple drivers of the cumulative number of COVID-19 cases. Our regression model adequately explains the cumulative COVID-19 case numbers (per 1 million population). As the COVID-19 spread progressed, the explanatory power ( $R^2$ ) of the model increased, reaching > 70% in April 2020. Climate, host mobility, and host susceptibility to COVID-19 largely explained the variance among COVID-19 case numbers across locations; the relative importance of host mobility and that of host susceptibility to COVID-19 were both greater than that of climate. Notably, the relative importance of these factors changed over time; the number of days from outbreak onset drove COVID-19 spread in the early stage, then human mobility accelerated the pandemic, and lastly climate (temperature) propelled the phase following disease expansion. Our findings demonstrate that the COVID-19 pandemic is deterministically driven by climate suitability, cross-border human mobility, and region-specific COVID-19 susceptibility. The identification of these multiple drivers of the COVID-19 outbreak trajectory, based on mapping the spread of COVID-19, will contribute to a better understanding of the COVID-19 disease transmission risk and inform long-term preventative measures against this disease.

## Introduction

The spread of infectious diseases through host-pathogen interaction is fundamentally underpinned by macroecological and biogeographical processes [1, 2]; key processes include virus origination, dispersal, and evolutionary diversification through local transmissions in human societies [3]. Since December 2019, coronavirus disease 2019 (COVID-19), caused by sudden



## ECOLOGY

# Occurrence-based diversity estimation reveals macroecological and conservation knowledge gaps for global woody plants

Buntarou Kusumoto<sup>1,2,3,4,5\*†</sup>, Anne Chao<sup>6</sup>, Wolf L. Eiserhardt<sup>5,7</sup>, Jens-Christian Svenning<sup>7,8</sup>, Takayuki Shono<sup>2,4</sup>, Yasuhiro Kubota<sup>2,4,9</sup>

Incomplete sampling of species' geographic distributions has challenged biogeographers for many years to precisely quantify global-scale biodiversity patterns. After correcting for the spatial inequality of sample completeness, we generated a global species diversity map for woody angiosperms (82,974 species, 13,959,780 occurrence records). The standardized diversity estimated more pronounced latitudinal and longitudinal diversity gradients than the raw data and improved the spatial prediction of diversity based on environmental factors. We identified areas with potentially high species richness and rarity that are poorly explored, unprotected, and threatened by increasing human pressure: They are distributed mostly at low latitudes across central South America, Central Africa, subtropical China, and Indomalayan islands. These priority areas for botanical exploration can help to efficiently fill spatial knowledge gaps for better describing the status of biodiversity and improve the effectiveness of the protected area network for global woody plant conservation.

Copyright © 2023 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

## INTRODUCTION

The accumulation of species occurrence data is a fundamental basis for biodiversity science, providing rising opportunities toward addressing major challenges in ecology and conservation (1, 2). Occurrence records have been widely used to model species distribution (3) and to estimate diversity at given localities (4). However, occurrence records notoriously suffer from incompleteness and biases (5), where observed species diversity is statistically influenced by sample size (6). As most occurrence records stem from collections taken for purposes other than estimating diversity patterns, their coverage is usually not geographically systematic nor comprehensive, resulting in a dominance of omission errors (7, 8); to complicate matters, species absence is scale-dependent, and its information is usually unavailable (9). This so-called Wallacean shortfall in biodiversity knowledge (10) potentially precludes a solid understanding of geographical biodiversity patterns (11, 12) and implementation of spatial conservation planning (13).

To correctly capture species diversity patterns, knowing the geographic variation in sample completeness of species occurrence data is critical (14). The explicit link between sample size, completeness, and diversity enables standardization of an observed diversity using rarefaction or extrapolation based on sample completeness (15). This allows fair comparisons of species diversity across multiple assemblages measured at unequal sample completeness without necessarily knowing their true diversity (14). Notably, latitudinal and

longitudinal diversity gradients have recently been revisited in this manner, especially in marine ecosystems (16, 17), and revealed unexpected diversity patterns (e.g., bi- or multimodality). Thus, diversity estimation theory challenges the generality of macroecological patterns that often suffer from serious sampling bias (18).

To achieve the global goals and milestones to counteract the current biodiversity crisis [e.g., the post-2020 Biodiversity Framework; (19)], the spatial allocation of conservation resources (e.g., land areas) is a key issue. For effective avoidance or mitigation of negative human impacts, spatial planning based on reliable information of biodiversity distribution is essential (20). However, spatial planning analyses implicitly assume that biodiversity patterns are accurately described, hereunder equally so inside and outside existing conservation areas; the validity of this assumption has not been examined at a global scale yet.

In this study, we focused on the species diversity of woody angiosperms. Woody angiosperms play a crucial role as ecosystem engineers, shaping most terrestrial biomes and supporting ecosystem functions and services on Earth (21). A recent study applied diversity estimation theory to a global occurrence record dataset and estimated the continental-level tree species richness, correcting for uneven sample completeness (4). However, their analysis did not include all woody angiosperms and only estimated diversity at the level of bioregions (biomes on continents). Global patterns of woody plant diversity at finer resolutions remain to be estimated from occurrence records and compared to previous studies using different data sources such as floristic checklists (22, 23) and plot surveys (24).

Here, we generated a global diversity map for woody angiosperms using 13,959,780 occurrence records for 82,974 species. We computed sample completeness and standardized species diversities using a Hill number-based approach to examine bias-corrected geographical patterns of species diversity. Hill numbers (or the effective number of species) (25) have been increasingly used to quantify the species diversity of assemblages. In particular, we

<sup>1</sup>Faculty of Agriculture, Kyushu University, Fukuoka, Japan. <sup>2</sup>Think Nature Inc., Naha City, Japan. <sup>3</sup>University Museum, University of the Ryukyus, Nishihara, Japan.

<sup>4</sup>Faculty of Science, University of the Ryukyus, Nishihara, Japan. <sup>5</sup>Royal Botanic Gardens, Kew, UK. <sup>6</sup>National Tsing Hua University, Hsinchu, Taiwan. <sup>7</sup>Section for Econinformatics and Biodiversity, Department of Biology, Aarhus University, Aarhus, Denmark. <sup>8</sup>Center for Ecological Dynamics in a Novel Biosphere (ECONOVE) and for Biodiversity Dynamics in a Changing World (BIOCHANGE), Department of Biology, Aarhus University, Aarhus, Denmark. <sup>9</sup>Tropical Biosphere Research Center, University of the Ryukyus, Nishihara, Japan.

\*Corresponding author. Email: kusumoto.buntarou@gmail.com

†Present address: Kasuya Research Forest, Kyushu University, 394 Tsubakuro, Sasaguri, Kasuya, Fukuoka, Japan.

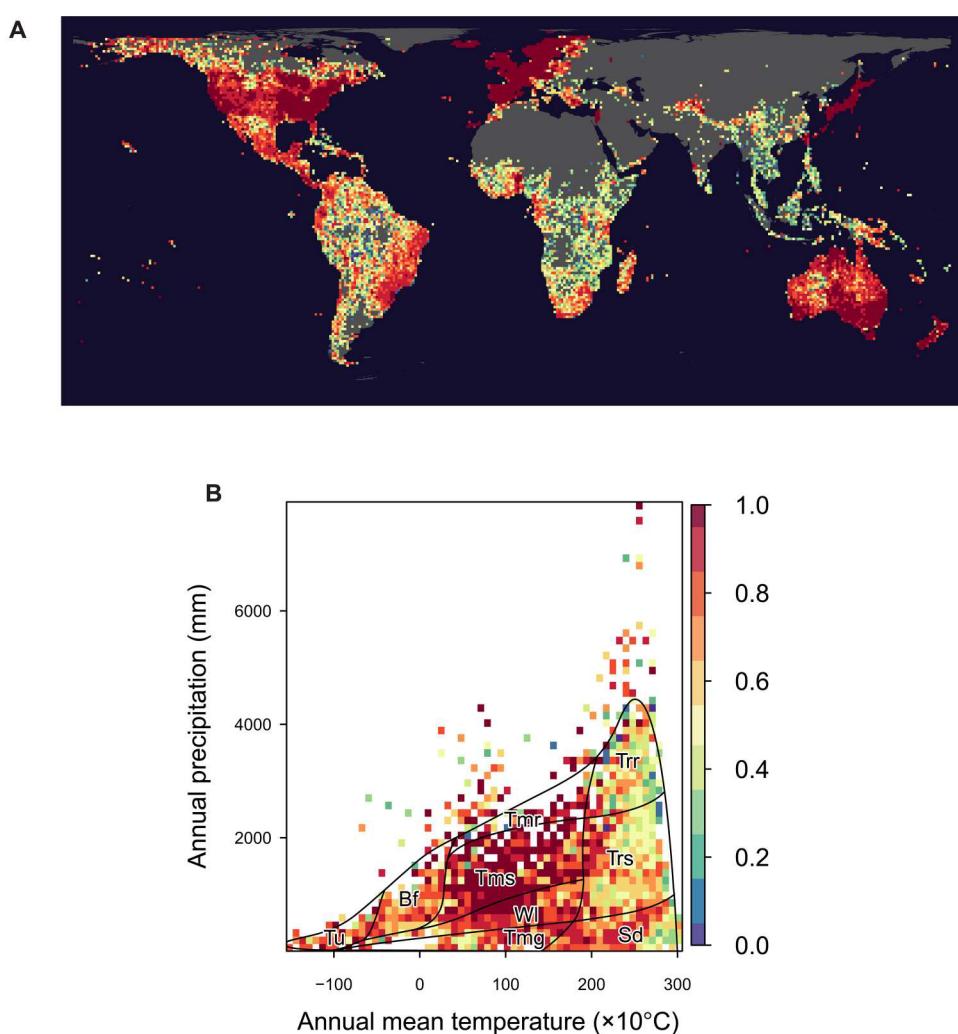
evaluated the impact of sample completeness on the description of species diversity and ecological inferences, especially of latitudinal and longitudinal diversity patterns related to spatial resolution, and identified predominant environmental drivers of species diversity at the global and regional scales. We also examined the spatial congruence of the species diversity and sample completeness with the global protected area network and changes in the human pressure. Last, we identified spatial priority areas for allocation of future sampling effort to effectively fill knowledge gaps.

## RESULTS AND DISCUSSION

### Observed diversity and sample completeness

Sample completeness measured by sample coverage, a concept originally developed by Alan Turing in his cryptographic analysis during World War II, greatly varied globally for the occurrence records of woody angiosperms (Fig. 1) (5, 26). Sample coverage is defined as the proportion of the total number of incidences

(counted by the 10 km-by-10 km subcells) belonging to detected species to the entire incidences including detected and undetected species. It tended to be high in temperate regions, including North America, Europe, Japan, Australia, and New Zealand. This trend likely reflects the sociopolitical histories of botanical collections rather than climatic conditions (27). Such geographical inequality of sampling effort distorts the description, interpretation, and prediction of biodiversity patterns (26, 28) because the observed diversity patterns reflect both multiple gradients of true diversity and the spatial bias of sampling efforts (29). The observed number of species showed a strong spatial congruence with the total number of occurrences (fig. S1). Expectedly, sample coverage was lower and more variable at the finest spatial resolution (100 km by 100 km) than at coarse resolution (~800 km by 800 km) (fig. S2). Such positive scale dependency of sample completeness has been reported previously in a regional-scale study of plants (30) and a global-scale study of stony corals (17).



**Fig. 1. Sample completeness (sample coverage) of species occurrence records of woody angiosperms at global scale.** (A) Geographical map at the 100 km-by-100 km equal-area grids ( $n = 8427$ ) and (B) the distribution on Whittaker's biome plot: tundra (Tu), boreal forest (Bf), temperate grassland/desert (Tmg), woodland/shrubland (WI), temperate seasonal forest (Tms), temperate rain forest (Tmr), tropical rain forest (Trr), tropical seasonal forest/savanna (Trs), and subtropical desert (Sd). In (B), the sample coverage values were aggregated to the median values in pixels divided  $60 \times 60$  of the climate space.



# Past and future decline of tropical pelagic biodiversity

Moriaki Yasuhara (安原盛明)<sup>a,b,1,2</sup>, Chih-Lin Wei<sup>c,1</sup>, Michal Kucera<sup>d,e</sup>, Mark J. Costello<sup>f,g</sup>, Derek P. Tittensor<sup>h,i</sup>, Wolfgang Kiessling<sup>j</sup>, Timothy C. Bonebrake<sup>b</sup>, Clay R. Tabor<sup>k</sup>, Ran Feng<sup>k</sup>, Andrés Baselga<sup>l,m</sup>, Kerstin Kretschmer<sup>d,e</sup>, Buntarou Kusumoto<sup>n</sup>, and Yasuhiro Kubota (久保田康裕)<sup>n</sup>

<sup>a</sup>Swire Institute of Marine Science, The University of Hong Kong, Hong Kong Special Administrative Region, China; <sup>b</sup>School of Biological Sciences, The University of Hong Kong, Hong Kong Special Administrative Region, China; <sup>c</sup>Institute of Oceanography, National Taiwan University, 106 Taipei, Taiwan; <sup>d</sup>MARUM–Center for Marine Environmental Sciences, University of Bremen, 28359 Bremen, Germany; <sup>e</sup>Faculty of Geosciences, University of Bremen, 28359 Bremen, Germany; <sup>f</sup>School of Environment, The University of Auckland, 1142 Auckland, New Zealand; <sup>g</sup>Faculty of Biosciences and Aquaculture, Nord University, 8049 Bodø, Norway; <sup>h</sup>Department of Biology, Dalhousie University, Halifax, NS, B3H 4R2 Canada; <sup>i</sup>United Nations Environment Programme World Conservation Monitoring Centre, CB3 0DL Cambridge, United Kingdom; <sup>j</sup>GeoZentrum Nordbayern, Department of Geography and Geosciences, Friedrich-Alexander Universität Erlangen–Nürnberg, 91054 Erlangen, Germany; <sup>k</sup>Department of Geosciences, University of Connecticut, Storrs, CT 06269; <sup>l</sup>Departamento de Zoología, Genética y Antropología Física, Facultad de Biología, Universidad de Santiago de Compostela, 15782 Santiago de Compostela, Spain; <sup>m</sup>CRETUS Institute, Universidad de Santiago de Compostela, 15782 Santiago de Compostela, Spain; and <sup>n</sup>Faculty of Science, University of the Ryukyus, 903-0213 Okinawa, Japan

Edited by James P. Kennett, University of California, Santa Barbara, CA, and approved April 15, 2020 (received for review October 1, 2019)

A major research question concerning global pelagic biodiversity remains unanswered: when did the apparent tropical biodiversity depression (i.e., bimodality of latitudinal diversity gradient [LDG]) begin? The bimodal LDG may be a consequence of recent ocean warming or of deep-time evolutionary speciation and extinction processes. Using rich fossil datasets of planktonic foraminifera, we show here that a unimodal (or only weakly bimodal) diversity gradient, with a plateau in the tropics, occurred during the last ice age and has since then developed into a bimodal gradient through species distribution shifts driven by postglacial ocean warming. The bimodal LDG likely emerged before the Anthropocene and industrialization, and perhaps ~15,000 y ago, indicating a strong environmental control of tropical diversity even before the start of anthropogenic warming. However, our model projections suggest that future anthropogenic warming further diminishes tropical pelagic diversity to a level not seen in millions of years.

latitudinal diversity gradients | planktonic foraminifera | temperature | Last Glacial Maximum | climate change

Latitudinal diversity gradients (LDGs), the equatorially centered parabolic diversity patterns, have been described for over 200 y in terrestrial systems (1–4) and are also well established in marine environments (5–7). However, there is an increasing recognition that marine LDGs, particularly those in open-ocean systems, tend to have a tropical diversity depression and thus, to be bimodal (8–14).

This current tropical depression is consistent with present-day temperatures being beyond the upper physiological thermal tolerances of some species. An inability of species to tolerate high temperatures or sustained physiological stresses may cause shifts of their latitudinal ranges farther poleward as the climate warms. Indeed, a near-future tropical biodiversity decline has been predicted with ongoing human-induced climate warming (15–19), and ecosystem-scale impacts of ocean warming are already evident (20–24).

Alternatively, or additionally, the current tropical dip in diversity could be explained through an evolutionary mechanism of higher speciation rates and/or lower extinction rates at the edges of the tropics (8, 13). Distinguishing the ecological and evolutionary timescale processes responsible for observed variations in the shape of marine LDGs is critical for assessing the outcome of biotic responses to rapid anthropogenic warming over the coming century (12). However, the lack of a standardized paleoecological baseline for the pelagic LDG has compromised separating whether the observed bimodality is caused by a rapid ecological response to ocean warming, by a longer-term and slower evolutionary process, or both (e.g., ref. 14). While several paleontological studies have shown bimodal LDGs (25), they are not directly comparable with the present-day pelagic bimodality or

do not answer this question directly, because they are terrestrial, not global in extent, or too deep time (e.g., Paleogene or Mesozoic) to evaluate the hypothesis of rapid ecological response.

The calcified shells of planktonic foraminifera, abundant and widespread protists in the world's oceans, are well preserved in marine sediments and can thus provide a baseline for tracking trends in the LDG over the geologic past (26, 27). In addition, the relationship between temperature and planktonic foraminiferal diversity is consistent with that of many other open-ocean organisms (5, 11, 28). Here, we use global datasets of pre-industrial (broadly representing a Late Holocene situation) (*Materials and Methods*) and Last Glacial Maximum (LGM; ca. 21 ky ago) planktonic foraminifera as well as a future diversity projection to provide empirical evidence that the tropical diversity depression is neither a recent anthropogenic phenomenon nor of deep-time origin. Rather, it was likely caused by a post–ice-age warming, suggesting a major role for distributional shifts driven by climate.

## Results and Discussion

**Diversity Patterns with Latitude and Temperature.** Our global analysis of planktonic foraminiferal diversity (calculated as species richness [Hill number,  $q = 0$ ] and effective number of common species [Hill number,  $q = 1$ ]) (*Materials and Methods*)

## Significance

We discovered that the tropical oceanic diversity depression is not a recent phenomenon nor very deep time in origin by using a comprehensive global dataset of the calcified shells of planktonic foraminifera, abundant unicellular organisms in the world's oceans, which are exceptionally well preserved in marine sediments as fossils. The diversity decline in the lowest latitudes may have started due to rapid post–ice-age warming around 15,000 y ago. Warming may by the end of this century diminish tropical oceanic diversity to an unprecedented level in human history.

Author contributions: M.Y. designed research; M.Y., C.-L.W., M.K., M.J.C., D.P.T., W.K., T.C.B., C.T., R.F., A.B., K.K., B.K., and Y.K. performed research; M.Y., C.-L.W., and D.P.T. analyzed data; and M.Y., C.-L.W., M.K., M.J.C., D.P.T., W.K., T.C.B., C.T., R.F., A.B., K.K., B.K., and Y.K. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

<sup>1</sup>M.Y. and C.-L.W. contributed equally to this work.

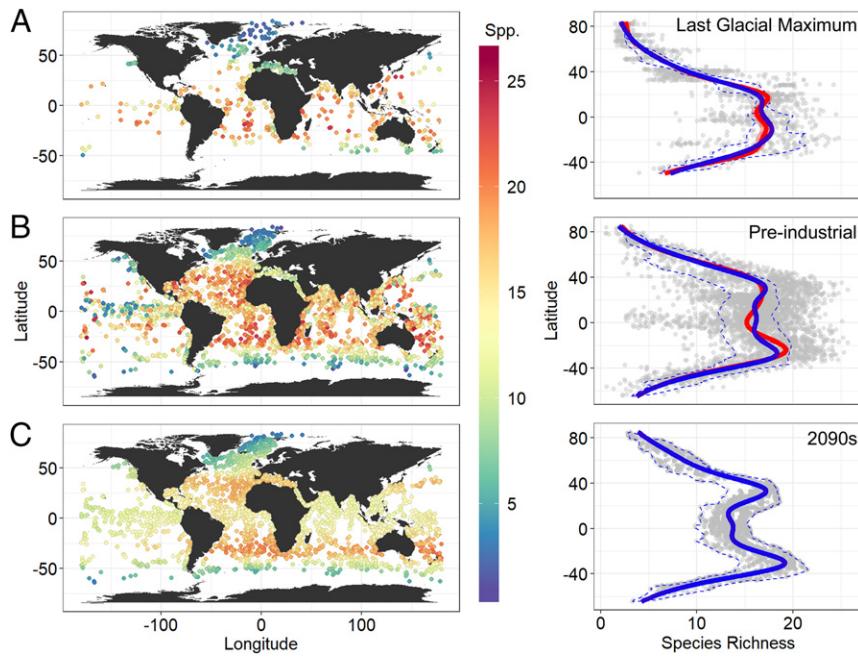
<sup>2</sup>To whom correspondence may be addressed. Email: moriakiyasuhara@gmail.com.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1916923117/DCSupplemental>.

First published May 26, 2020.

ECOLOGY

EARTH, ATMOSPHERIC,  
AND PLANETARY SCIENCES



**Fig. 1.** Species richness of planktonic foraminifers during the (A) LGM, (B) PIC, and for (C) 2091 to 2100 (2090s) as maps and latitudinal gradients. Colored and gray dots (in the maps and the latitudinal gradients, respectively) indicate the observed diversities in A (LGM) and B (PIC). These observed LGM and PIC diversities were modeled by SST, coordinates, and ocean basin using a GAM to predict the diversities in 2090s (colored and gray dots in C) with future SST (based on RCP 8.5) as well as those during the LGM and PIC themselves. The predicted latitudinal diversities for the three time periods (enclosed by blue dashed lines) were smoothed by a GAM to show LDGs (blue lines). The latitudinal gradients of observed diversities during the LGM and PIC were also fitted by a GAM and shown as the red lines with the shaded areas indicating the 95% CIs (the shaded area is small, overlaps the red line, and so is not visible in the PIC panel). For the LGM and PIC gray dots, a small amount of jitter was added on the x axis to make them visible when overlapping. *SI Appendix, Fig. S1* shows empirical and projected diversities using a Hill number of order  $q = 1$ .

demonstrates that during the LGM, the LDG was unimodal (or only weakly bimodal), whereas the preindustrial LDG was bimodal with a distinct tropical diversity depression (Fig. 1 and *SI Appendix, Fig. S1* and *Tables S1 and S2*). This indicates that the strength of the bimodal LDG for planktonic foraminifers cannot be entirely due to long-term evolutionary processes because it was minimal during the LGM (Fig. 1 and *SI Appendix, Fig. S1* and *Tables S1 and S2*), and there have been no known global extinctions or speciations of any planktonic foraminiferal species since the LGM (29).

We propose that the cause of the bimodality may then be environmentally driven extirpation and/or immigration. During warming, any diversity losses at higher latitudes (due to range shifts of species to even higher latitudes) are compensated for by the poleward movements of species from lower latitudes. However, in the tropics, such compensation due to species range shifts is not possible, resulting in a tropical diversity decline (15, 17, 30, 31).

It is unlikely that the tropical diversity depression is a very recent phenomenon originating in the Anthropocene because we found that the preindustrial LDG was already bimodal. Thus, the bimodal LDG most likely developed during the post-LGM warming, with a 5.2% loss in the mean projected species richness since the LGM at the equator (calculated based on the mean predictions within  $\pm 1^\circ$  latitude) (Fig. 1).

The LDG exhibited a tropical plateau (or weak bimodality) during the LGM (Fig. 1 and *SI Appendix, Fig. S1*) indicating an approach toward diversity saturation (at or beyond the optimum in the unimodal temperature–diversity relationship; see the next paragraph) with relatively low maximum global sea temperature. The distinct tropical diversity decline may have begun  $\sim 15,000$  y ago, given that a rapid postglacial warming started at that time (32). The duration of glacial periods has been much longer than that of interglacial periods during the Late Quaternary. Therefore, the tropical thermal niches of marine organisms may be optimized

to the maximum temperatures of glacial periods, leading to tropical diversity depressions during warm periods, given that marine niche conservatism is known to have existed during Late Quaternary climate changes (33). As a bimodal LDG is known to be present during the last interglacial (in corals) (34), it is likely that the bimodal LDG has appeared repeatedly during warm interglacial periods during the Late Quaternary and weakened during glacial periods. Species adapted to very warm temperatures existed during the Pliocene, the major previous warmer-than-present period, but significant extinctions of these species are known during the Plio–Pleistocene cooling (27). Note that pre-Plio–Pleistocene Phanerozoic LDGs are also known to be dynamic (14, 35–37), although the underlying mechanism may be different.

Sea surface temperature (SST) has been and is unimodal with latitude (Fig. 2D) (the next paragraph discusses the equatorial upwelling zone). It is also predicted to remain unimodal under the RCP 8.5 “business-as-usual” climate warming scenario in 2091 to 2100 (“2090s” hereafter), with  $\sim 0$  to  $4^\circ\text{C}$  warming relative to the preindustrial control (PIC) (Fig. 2). The magnitude of the predicted warming from the PIC to the RCP 8.5 2090s will be larger (and much more rapid) than that from the LGM to PIC (Fig. 2), particularly in the tropics. The unimodal (or only weakly bimodal) LDG during the LGM and the bimodal LDG during the preindustrial time period reflect a positive temperature–diversity relationship from  $-2$  to  $20^\circ\text{C}$  and a negative relationship beyond that, especially beyond  $25^\circ\text{C}$  and for species richness (*SI Appendix, Fig. S2*). Thus, the present reduction of species diversity in the tropics is likely due to high sea temperatures (*SI Appendix, Fig. S2*), a thermal response also identified in other pelagic groups (38). Such very high mean temperatures (those exceeding  $25^\circ\text{C}$ ) did not exist in any latitudinal band during the LGM (Fig. 2). Supporting our interpretation is the observation that planktonic foraminifer species tend to have optimum temperature

**SPECIAL FEATURE**

Approaches for general rules of biodiversity patterns in space and time

# Quantifying sample completeness and comparing diversities among assemblages

Anne Chao<sup>1</sup> | Yasuhiro Kubota<sup>2</sup> | David Zelený<sup>3</sup> | Chun-Huo Chiu<sup>4</sup> |  
 Ching-Feng Li<sup>5</sup> | Buntarou Kusumoto<sup>2,6</sup> | Moriaki Yasuhara<sup>7</sup> |  
 Simon Thorn<sup>8</sup> | Chih-Lin Wei<sup>9</sup> | Mark J. Costello<sup>10,11</sup> | Robert K. Colwell<sup>12,13</sup>

<sup>1</sup>Institute of Statistics, National Tsing Hua University, Hsin Chu, Taiwan

<sup>2</sup>Faculty of Science, University of the Ryukyus, Nishihara, Japan

<sup>3</sup>Institute of Ecology and Evolutionary Biology, National Taiwan University, Taipei, Taiwan

<sup>4</sup>Department of Agronomy, National Taiwan University, Taipei, Taiwan

<sup>5</sup>School of Forestry and Resource Conservation, National Taiwan University, Taipei, Taiwan

<sup>6</sup>Royal Botanic Gardens, Kew, Richmond, UK

<sup>7</sup>School of Biological Sciences and Swire Institute of Marine Science, The University of Hong Kong, Pokfulam Road, Hong Kong

<sup>8</sup>Field Station Fabrikschleichach, Biocenter, University of Würzburg, Rauhenebrach, Germany

<sup>9</sup>Institute of Oceanography, National Taiwan University, Taipei, Taiwan

<sup>10</sup>School of Environment, University of Auckland, Auckland, New Zealand

<sup>11</sup>Nord University, Bodø, Norway

<sup>12</sup>Department of Ecology and Evolution, University of Connecticut, Storrs, USA

<sup>13</sup>University of Colorado Museum of Natural History, Boulder, USA

## Correspondence

Anne Chao, Institute of Statistics,  
 National Tsing Hua University, Hsin Chu  
 30043, Taiwan.  
 Email: chao@stat.nthu.edu.tw

## Funding information

Taiwan Ministry of Science and  
 Technology, Grant/Award Number:  
 107-2118-M007-006-MY3; Research Grants  
 Council of Honk Kong, Grant/Award  
 Numbers: HKU 709413P, HKU 17303115,  
 HKU 173025518; University of Hong  
 Kong, Grant/Award Number:  
 201111159140, 201611159053,  
 201711159057; JSPS KAKENHI, Grant/  
 Award Number: 16H02755; University of  
 the Ryukyus

## Abstract

We develop a novel class of measures to quantify sample completeness of a biological survey. The class of measures is parameterized by an order  $q \geq 0$  to control for sensitivity to species relative abundances. When  $q = 0$ , species abundances are disregarded and our measure reduces to the conventional measure of completeness, that is, the ratio of the observed species richness to the true richness (observed plus undetected). When  $q = 1$ , our measure reduces to the sample coverage (the proportion of the total number of individuals in the entire assemblage that belongs to detected species), a concept developed by Alan Turing in his cryptographic analysis. The sample completeness of a general order  $q \geq 0$  extends Turing's sample coverage and quantifies the proportion of the assemblage's individuals belonging to detected species, with each individual being proportionally weighted by the  $(q - 1)$ th power of its abundance. We propose the use of a continuous profile depicting our proposed

This paper is dedicated to the memory of our coauthor Ching-Feng (Woody) Li, who passed away on November 29, 2019, after a courageous battle with lung cancer. We hope Woody is free, without pain, happily hiking somewhere around Chilai mountain in Taiwan.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Ecological Research* published by John Wiley & Sons Australia, Ltd on behalf of The Ecological Society of Japan

**TABLE 1** Theoretical formulas and analytic estimators for sample completeness measures of order  $q \geq 0$  and three special cases ( $q = 0$ , 1 and 2) for (a) abundance data and (b) sampling-unit-based replicated incidence data. See the text and footnotes for notation and details

Order $q$	Theoretical formula	Analytic estimator
(a) Abundance data		
$q \geq 0$	${}^q C = \frac{{}^q \hat{\lambda}_{\text{detected}}}{{}^q \hat{\lambda}} = \frac{\sum_{i \in \text{detected}} p_i^q}{\sum_{i=1}^S p_i^q}$	$\hat{C} = \frac{{}^q \hat{\lambda}_{\text{detected}}}{{}^q \hat{\lambda}} = 1 - \frac{f_1}{n} \left[ \frac{A^{q-1}(1-A)}{{}^q \hat{\lambda}} \right]$
$q = 0$	${}^0 C = \frac{S_{\text{obs}}}{S}$	$\hat{C} = \frac{S_{\text{obs}}}{S_{\text{Chao1}}}$
$q = 1$	${}^1 C = \sum_{i \in \text{detected}} p_i$	$\hat{C} = 1 - \frac{f_1}{n}(1-A)$
$q = 2$	${}^2 C = \frac{\sum_{i \in \text{detected}} p_i^2}{\sum_{i=1}^S p_i^2}$	$\hat{C} = 1 - \frac{f_1}{n} \left[ \frac{A(1-A)}{\sum_{X_i \geq 2} X_i(X_i-1)/[n(n-1)]} \right]$
(b) Replicated incidence data		
$q \geq 0$	${}^q \Phi = \frac{{}^q \hat{\phi}_{\text{detected}}}{{}^q \hat{\phi}} = \frac{\sum_{i \in \text{detected}} \pi_i^q}{\sum_{i=1}^S \pi_i^q}$	$\hat{C} = \frac{{}^q \hat{\phi}_{\text{detected}}}{{}^q \hat{\phi}} = 1 - \frac{Q_1}{T} \left[ \frac{B^{q-1}(1-B)}{{}^q \hat{\phi}} \right]$
$q = 0$	${}^0 \Phi = \frac{S_{\text{obs}}}{S}$	$\hat{C} = \frac{S_{\text{obs}}}{S_{\text{Chao2}}}$
$q = 1$	${}^1 \Phi = \frac{\sum_{i \in \text{detected}} \pi_i}{\sum_{i=1}^S \pi_i}$	$\hat{C} = 1 - \frac{Q_1}{T} \frac{(1-B)}{\sum_{Y_i \geq 1} Y_i/T} = 1 - \frac{Q_1}{U}(1-B)$
$q = 2$	${}^2 \Phi = \frac{\sum_{i \in \text{detected}} \pi_i^2}{\sum_{i=1}^S \pi_i^2}$	$\hat{C} = 1 - \frac{Q_1}{T} \left[ \frac{B(1-B)}{\sum_{Y_i \geq 2} Y_i(Y_i-1)/[T(T-1)]} \right]$

Note: (1) See the Supporting Information for the formulas of  ${}^q \hat{\lambda}$  (Equation S1.6),  ${}^q \hat{\lambda}_{\text{detected}}$  (Equation S1.9),  ${}^q \hat{\phi}$  (Equation S2.6), and  ${}^q \hat{\phi}_{\text{detected}}$  (Equation S2.9), where  ${}^q \lambda = \sum_{i=1}^S p_i^q$  and  ${}^q \Phi = \sum_{i=1}^S \pi_i^q$  denote, respectively, the  $q$ th power sum for abundance and incidence data.

(2)  $A$  (for abundance data): the estimated mean relative frequency of singletons;  $B$  (for incidence data): the estimated mean detection probability in any sampling unit of unique species.

$$A = \begin{cases} 2f_2/[(n-1)f_1 + 2f_2], & \text{if } f_2 > 0; \\ 2/[(n-1)(f_1-1) + 2], & \text{if } f_2 = 0, f_1 \neq 0; \\ 1, & \text{if } f_2 = f_1 = 0. \end{cases} \quad B = \begin{cases} 2Q_2/[(T-1)Q_1 + 2Q_2], & \text{if } Q_2 > 0; \\ 2/[(T-1)(Q_1-1) + 2], & \text{if } Q_2 = 0, Q_1 > 0; \\ 1, & \text{if } Q_1 = Q_2 = 0. \end{cases}$$

(3)  $U = \sum_{Y_i \geq 1} Y_i$  denotes the total number of incidences based on detection/non-detection records of  $T$  sampling units.

individual is weighted by  $1/p$ , the inverse of its species relative abundance. Therefore, this measure is disproportionately sensitive to rare species, compared to measures with order  $q > 0$ .

(2) When  $q = 1$ , the measure  ${}^1 C$  reduces to the sum of the relative abundances of the detected species, or, equivalently, the fraction of the assemblage's individuals that belong to the detected species. This is the concept of Turing's sample coverage (Good, 1953, 2000), which quantifies sample completeness when all individuals are treated equally. The weight for every individual is the same, regardless of species, so that a species' weight is proportional to its abundance, without disproportionately favoring either abundant or rare species.

(3) When  $q = 2$ , the measure quantifies the fraction of the total number of individuals in the assemblage that belong to the detected species, with each species being proportionally weighted by its squared species relative abundance, or, equivalently, with each individual being proportionally weighted by its species abundance. Thus, the measure  ${}^2 C$  is disproportionately sensitive to highly abundant species. In

most surveys, highly abundant species would be detected in any sample; thus, the second- and higher-order measures typically yield values very close to unity.

The sample completeness measure of any order  $q \geq 0$  quantifies a generalized sample coverage, that is, the proportion of the total number of individuals in the assemblage belonging to detected species, with each species being proportionally weighted by  $p^q$ , the  $q$ th power of its species abundance. Equivalently, each individual is proportionally weighted by  $p^{q-1}$ . Our measures of orders  $q > 1$  are disproportionately sensitive to highly abundant species, whereas the measures of orders  $q < 1$  are disproportionately sensitive to rare species. The measure of order  $q = 1$  reduces to Turing's sample coverage, as described above.

Although  $(p_1, p_2, \dots, p_S)$  is modeled in our theory as species relative abundances, our derivation is also valid under a more general model in which  $(p_1, p_2, \dots, p_S)$  represent species detection probabilities. Generally, the detection probability for any individual is a combination of species abundance and the individual's detectability,



# Reconciling Darwin's naturalization and pre-adaptation hypotheses: An inference from phylogenetic fields of exotic plants in Japan

Buntarou Kusumoto<sup>1,2</sup> | Fabricio Villalobos<sup>3</sup> | Takayuki Shiono<sup>2</sup> |  
 Yasuhiro Kubota<sup>2,4</sup>

<sup>1</sup>Royal Botanic Gardens Kew, Richmond, UK

<sup>2</sup>Faculty of Science, University of the Ryukyus, Nishihara, Japan

<sup>3</sup>Laboratorio de Macroecología Evolutiva, Red de Biología Evolutiva, Instituto de Ecología, A.C., Xalapa, Mexico

<sup>4</sup>Tropical Biosphere Research Center, University of the Ryukyus, Nishihara, Japan

## Correspondence

Buntarou Kusumoto, Royal Botanic Gardens Kew, Richmond, Surrey TW9 3AB, UK.  
 Email: kusumoto.buntarou@gmail.com

## Funding information

The Ministry of the Environment, Japan, Grant/Award Number: 4-1501 and 4-1802; University of the Ryukyus; Japan Society for the Promotion of Science

Handling Editor: Dr. Daniel Chapman

## Abstract

**Aim:** Understanding the causes and consequences of biological invasions remains a challenge for several disciplines, including biogeography. One major issue in overcoming this challenge is disentangling the confounding mechanisms of species invasiveness and community invasibility. Here, we tackle this issue by applying a novel approach based on the phylogenetic affinities between exotic species and natives in the recipient community to elucidate naturalization and pre-adaptation processes.

**Location:** Japan.

**Taxon:** Seed plants.

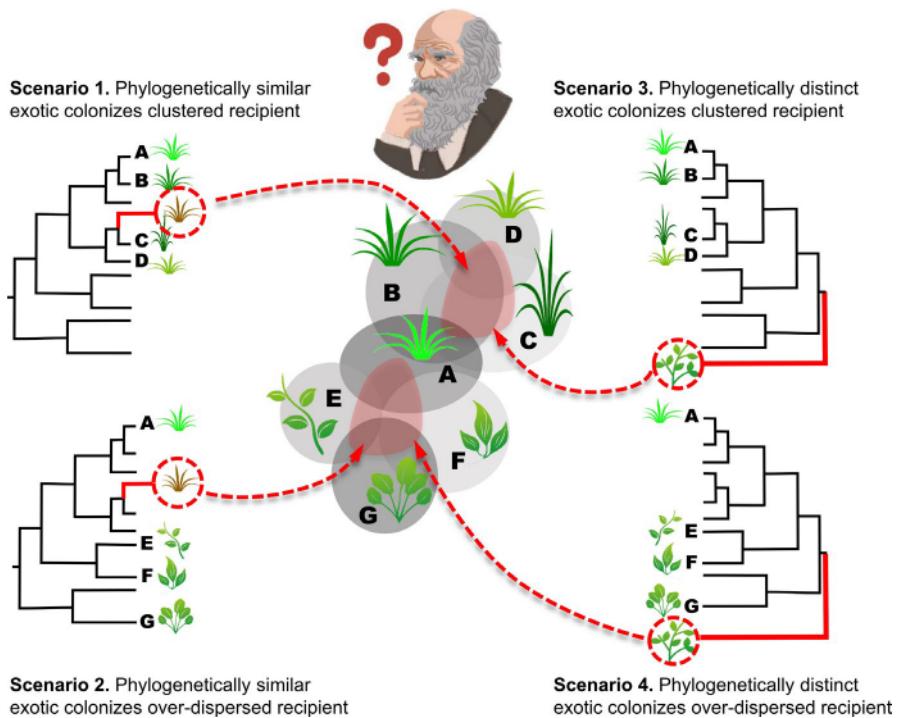
**Methods:** Geographical co-occurrence data for 1,094 exotics and 4,869 native species (including 1,676 endemics) were created at the 10-km grid-cell and vegetation-plot levels. For individual exotic species, standardized effect size of phylogenetic species variability ( $PSV_{SES}$ ) of the recipient native assemblage (i.e. phylogenetic fields) was calculated and its clustering/over-dispersion was tested, representing exotic invasiveness in relation to invasibility of native recipients. To identify drivers of species invasiveness, the correlation of  $PSV_{SES}$  with species attributes, involving phylogenetic distance between each exotic and native species, was explored.

**Results:** Phylogenetic fields ( $PSV_{SES}$ ) showed significant over-dispersion (~16% exotics) or clustering (~14% exotics). Interspecific variation of  $PSV_{SES}$  among exotics was substantially explained by species ecological attributes. Geographical extent and climatic niche widths were negatively correlated with  $PSV_{SES}$ . Preference for human influence was positively correlated with  $PSV_{SES}$  at the 10-km grid-cell level, but negatively at the vegetation-plot level. Exotics colonized from the Palearctic and Indo-Malay regions, which belong to the same biogeographical region as East Asia, tended to have clustered phylogenetic fields.

**Main conclusions:** Environmental filtering and biotic sorting both played a key role in exotic plant colonization, supporting both of Darwin's contradictory hypotheses of naturalization versus pre-adaptation. Clustered phylogenetic fields indicated that an exotic colonizes its recipient assemblage through abiotic filtering (i.e. pre-adaptation); at the same time, phylogenetic over-dispersion was indicative of naturalization for exotics that occupied a biotic niche space among native recipients (i.e. naturalization).



**FIGURE 1** Conceptual framework of the phylogenetic field approach for investigation of the relationship between exotic and recipient native plant species. There are four potential scenarios, which reflect the Darwin's conundrum, in the relationship between the phylogenetic structure in a recipient assemblage and phylogenetic relatedness of an exotic species with its recipient. Grey circles represent distribution ranges of native species (A–G). Red shaded areas represent the colonized area of an exotic species



Phylogenetic information has been used to describe both species invasiveness (Ricotta et al., 2010) and site invasibility (Davis, Grime, & Thompson, 2000). These concepts are interrelated and thus should be merged (Pyšek & Richardson, 2006). In this sense, the phylogenetic field approach can characterize the invasibility of a recipient assemblage for individual exotic species. This measure can substitute for a classical measure of species invasiveness, the phylogenetic distance between an exotic species and the recipients, which summarizes recipient information into a scalar (e.g. mean or minimum) (Ricotta et al., 2010). Hence, four possible scenarios are predicted (Figure 1): (a) a phylogenetically closely related exotic colonizes phylogenetically clustered recipients; (b) a phylogenetically closely related exotic colonizes phylogenetically over-dispersed recipients; (c) a phylogenetically distantly related exotic colonizes phylogenetically clustered recipients; and (d) a phylogenetically distantly related exotic colonizes phylogenetically over-dispersed recipients. Importantly, species-specific invasiveness should be revealed in conjunction with the invasibility of native recipients, as examined in site-based approaches (Gallien & Carboni, 2016; Lososová et al., 2015; Ng, Weaver, & Laport, 2018).

In the present study, we focused on exotic plant species naturalized in the East Asian islands across Japan and applied the phylogenetic field approach—characterizing each individual exotic species using the phylogenetic structure of all co-occurring native species across Japan, rather than within a single local site—to investigate species-specific invasiveness measured by invasibility of a recipient native assemblage, in the context of the naturalization and pre-adaptation processes. First, we compiled a species list of exotic seed plants that had colonized in Japan, and collected occurrence information for the exotic species from herbarium specimen records and vegetation census data. We then created a dataset of the

geographical co-occurrence for 1,094 exotic and 4,869 native seed plants species (Kubota, Shiono, & Kusumoto, 2015), and calculated the phylogenetic field of individual exotic species using their co-occurring recipient assemblage. To detect predominant drivers of invasiveness for exotics, we analysed the correlation of species-specific phylogenetic fields with species attributes. We also examined the association between phylogenetic fields of exotics and the phylogenetic distance (dissimilarity) between individual exotics and recipient natives. In addition, we repeated the same analyses at two spatial grains, namely 10-km grid cells and 1–5,000 m<sup>2</sup> vegetation plots, to capture macro-scale processes and local ecological processes, respectively, with two definitions of recipient assemblage (including all native and endemic species), and evaluated the potential impact of spatial resolution on co-occurrence among exotics and natives.

The phylogenetic field framework enables exploration of novel diversity patterns that characterize exotics based on their co-occurring native species. Phylogenetic field patterns of individual exotics are expected to reflect the ecological properties and/or introduction history of a species, which are assumed to be drivers of naturalization/invasion success. Therefore, we predicted that niche width and/or geographical distribution of exotic species may lead to both phylogenetic clustering and over-dispersion: on one hand, exotics with narrow niches could colonize a specific habitat in which the recipient native assemblage is likely to be phylogenetically clustered through environmental filtering (e.g. Kembel & Hubbell, 2006); on the other hand, exotics with broad niches could invade niche space among phylogenetically over-dispersed recipient natives, regardless of the habitat. Human disturbances, which release niche space or act as an abiotic filter for colonization of exotics (Jauni, Gripenberg, & Ramula, 2015; MacDougall & Turkington, 2005), could also shape the phylogenetic field of exotics by two mechanisms: colonization



## BIODIVERSITY LOSS

# The global loss of avian functional and phylogenetic diversity from anthropogenic extinctions

Thomas J. Matthews<sup>1,2\*</sup>, Kostas A. Triantis<sup>3</sup>, Joseph P. Wayman<sup>1</sup>, Thomas E. Martin<sup>4,5</sup>, Julian P. Hume<sup>6</sup>, Pedro Cardoso<sup>7,8</sup>, Søren Faurby<sup>9,10</sup>, Chase D. Mendenhall<sup>11</sup>, Paul Dufour<sup>12,13</sup>, François Rigal<sup>14</sup>, Rob Cooke<sup>15</sup>, Robert J. Whittaker<sup>16,17</sup>, Alex L. Pigot<sup>18</sup>, Christophe Thébaud<sup>19</sup>, Maria Wagner Jørgensen<sup>1</sup>, Eva Benavides<sup>1</sup>, Filipa C. Soares<sup>20</sup>, Werner Ulrich<sup>21</sup>, Yasuhiro Kubota<sup>22</sup>, Jon P. Sadler<sup>1</sup>, Joseph A. Tobias<sup>23†</sup>, Ferran Sayol<sup>24†</sup>

Humans have been driving a global erosion of species richness for millennia, but the consequences of past extinctions for other dimensions of biodiversity—functional and phylogenetic diversity—are poorly understood. In this work, we show that, since the Late Pleistocene, the extinction of 610 bird species has caused a disproportionate loss of the global avian functional space along with ~3 billion years of unique evolutionary history. For island endemics, proportional losses have been even greater. Projected future extinctions of more than 1000 species over the next two centuries will incur further substantial reductions in functional and phylogenetic diversity. These results highlight the severe consequences of the ongoing biodiversity crisis and the urgent need to identify the ecological functions being lost through extinction.

**T**he past 130,000 years have been characterized by substantial global environmental change as a result of natural climatic fluctuations and, increasingly, human actions, through drivers including habitat loss, hunting, introduced species, intensive agriculture, and climate change (1, 2). Anthropogenic drivers are known to have increased species extinction rates by orders of magnitude compared with the background extinction rate (1, 3, 4). Species losses have been especially severe on islands, with insular species representing ~75% of International Union for Conservation of Nature (IUCN)-documented post-1500 CE extinctions despite islands making up only ~7% of Earth's land area (2, 5).

Birds have been particularly affected, with hundreds of known extinctions (6–10). However, biodiversity is multidimensional, and the ecological and evolutionary consequences of this species loss are still not fully understood (11, 12). Birds contribute a range of important ecological functions, including pollination, predator-prey interactions, and seed dispersal

(13–17). The ecological roles of particular species are dictated by their functional traits—the morphological and ecological characteristics that determine an organism's fitness or performance (17–19). Thus, estimates of functional diversity (FD)—the range of functional traits of all species in an assemblage—can provide a more mechanistic understanding of the effects of extinctions on ecosystem function compared with the traditional focus on species richness (17, 19, 20). Additionally, phylogenetic diversity (PD)—the breadth of evolutionary history represented by a set of species—provides a complementary metric of ecological structure, offering insight into both the evolutionary processes shaping biodiversity and unmeasured niche dimensions that may not be captured in a given trait dataset (21–25). A combination of FD and PD therefore provides a vital window onto the ecological implications of extinction and the uniqueness of the species that have been lost.

Bird extinctions during the Late Pleistocene and Holocene, which on some archipelagos

represent most of the native avifauna (1), are thought to have reduced avian FD and PD (8), but to what extent is unclear. Given the apparent high functional overlap among bird species at global scales, a null expectation would be that anthropogenic extinctions have resulted in relatively small reductions in global FD and PD (16, 27). However, species traits are known to have influenced the susceptibility of island birds to extinction drivers (2, 10, 28). Hence, we may expect the loss of FD over this period to have exceeded that predicted by a null model that assumes no association between traits and extinction. If these traits are nonrandomly associated with phylogenetic uniqueness, we may also expect PD loss to have been greater than expected. To date, these combined hypotheses remain untested at the global scale.

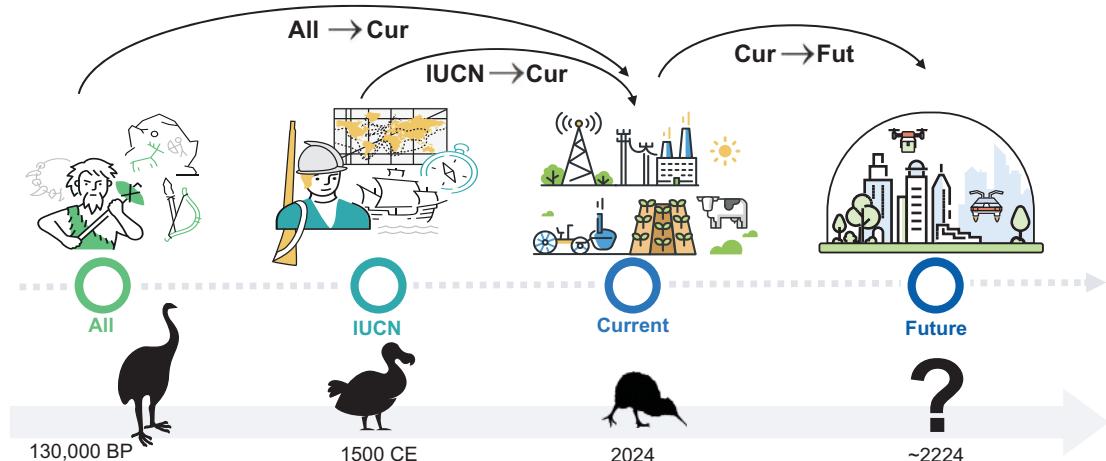
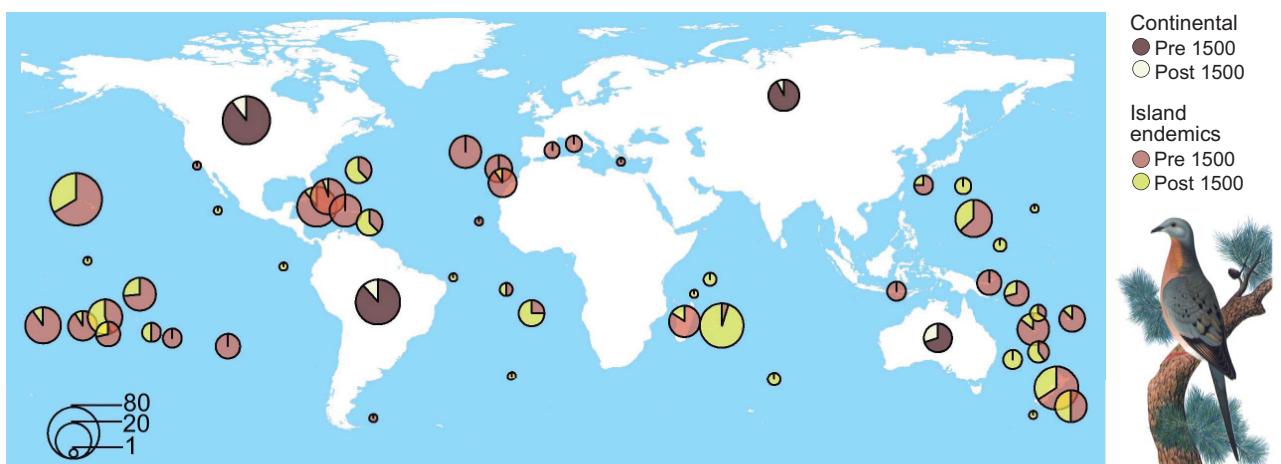
In this work, we provide complete global estimates of the avian FD and PD lost through anthropogenic extinctions over the past 130,000 years as well as estimates of the magnitude of expected future loss. As a first step, we compiled the most comprehensive dataset to date of all known bird extinctions during the Late Pleistocene and Holocene, distinguishing between anthropogenic extinctions and extinction events of unknown cause (29). For each extinct species, we measured eight functional traits (including beak, tarsus, and wing length) from museum skins and skeletal specimens (fig. S1). All are continuous traits previously shown to provide accurate and fine-grained information on the functional, behavioral, and trophic niches of birds (16, 27). To augment these measurements, we obtained published trait values from the literature where possible (including body mass) and filled remaining data gaps using Bayesian hierarchical probabilistic matrix factorization (29, 30). This dataset was combined with a dataset of traits measured using the same methods from all the world's 11,003 extant bird species (17).

Using these global datasets, we calculated the amount of avian FD that has been lost

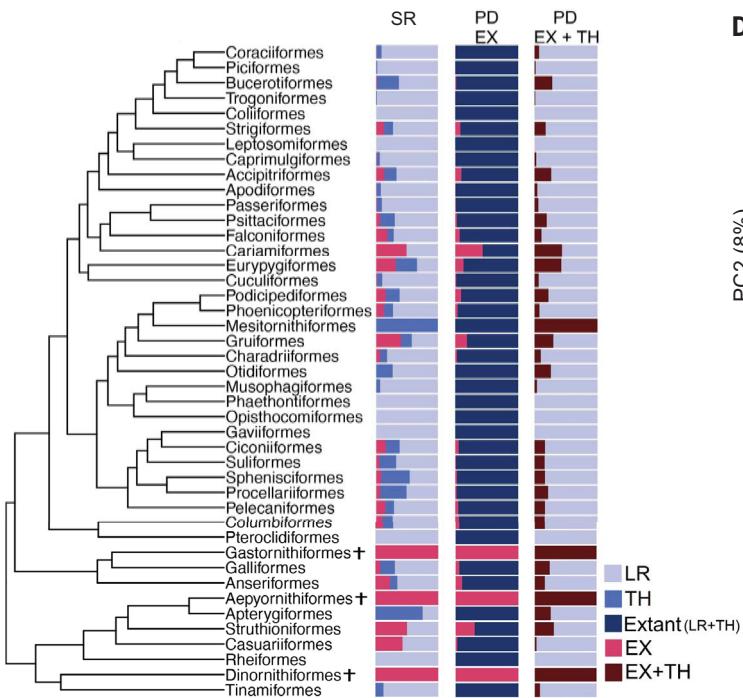
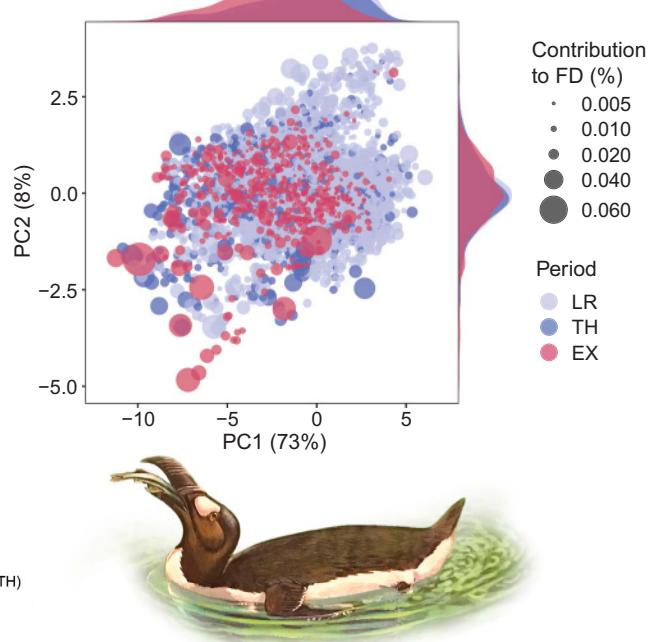
<sup>1</sup>School of Geography, Earth and Environmental Sciences (GEES) and Birmingham Institute of Forest Research, University of Birmingham, Birmingham, UK. <sup>2</sup>Centre for Ecology, Evolution and Environmental Changes (CE3C), Azorean Biodiversity Group, CHANGE – Global Change and Sustainability Institute, and Faculty of Agricultural Sciences and Environment, Universidade dos Açores, Angra do Heroísmo, Açores, Portugal. <sup>3</sup>Department of Ecology and Taxonomy, Faculty of Biology, National and Kapodistrian University of Athens, Athens GR-15784, Greece. <sup>4</sup>School of Natural Sciences, College of Environmental Sciences and Engineering, Bangor University, Bangor, UK. <sup>5</sup>Operation Wallacea, Wallace House, Old Bolingbroke, Lincolnshire, UK. <sup>6</sup>Bird Group, Life Sciences, Natural History Museum, Tring, UK. <sup>7</sup>Laboratory for Integrative Biodiversity Research (LIBRE), Finnish Museum of Natural History Luomus, University of Helsinki, Helsinki, Finland. <sup>8</sup>CE3C, CHANGE – Global Change and Sustainability Institute, Faculty of Sciences, University of Lisbon, Lisbon, Portugal. <sup>9</sup>Department of Biological and Environmental Sciences, University of Gothenburg, Gothenburg, Sweden. <sup>10</sup>Gothenburg Global Biodiversity Centre, Department of Biological and Environmental Sciences, University of Gothenburg, Gothenburg, Sweden. <sup>11</sup>Physician Assistant Studies, Slippery Rock University, Slippery Rock, PA 16057, USA. <sup>12</sup>Center for Functional and Evolutionary Ecology (CEFE), Université de Montpellier, CNRS, EPHE-PSL University, IRD, Montpellier, France. <sup>13</sup>Station Biologique de la Tour du Valat, Arles, France. <sup>14</sup>CNRS - Université de Pau et des Pays de l'Adour - E2S UPPA, Institut Des Sciences Analytiques et de Physico Chimie pour l'Environnement et les Matériaux, UMR5254, Pau, France. <sup>15</sup>UK Centre for Ecology & Hydrology, Crowmarsh Gifford, Wallingford, Oxfordshire, UK. <sup>16</sup>School of Geography and the Environment, University of Oxford, Oxford, UK. <sup>17</sup>Center for Macroecology, Evolution and Climate, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark. <sup>18</sup>Centre for Biodiversity and Environment Research, Department of Genetics, Evolution and Environment, University College London, London, UK. <sup>19</sup>Centre de Recherche sur la Biodiversité et l'Environnement (CRBE), UMR 5300 Centre National de la Recherche Scientifique (CNRS), Institut de Recherche pour le Développement (IRD), Université Paul Sabatier (Toulouse III), Toulouse Cedex 9, France. <sup>20</sup>CE3C, Departamento de Biología Animal, CHANGE – Global Change and Sustainability Institute, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal. <sup>21</sup>Department of Ecology and Biogeography, Nicolaus Copernicus University, Toruń, Poland. <sup>22</sup>Faculty of Science, University of the Ryukyus, Okinawa, Japan. <sup>23</sup>Department of Life Sciences, Imperial College London, Silwood Park, Ascot, UK. <sup>24</sup>CREAF, Edifici C Campus UAB, E08193 Cerdanyola del Vallès, Catalonia, Spain.

\*Corresponding author. Email: [tjm676@gmail.com](mailto:tjm676@gmail.com)

†These authors contributed equally to this work.

**A****B**

Downloaded from https://www.science.org at University of the Ryukyus on October 20, 2024

**C****D**

**Fig. 1. Overview of the study design and a summary of the FD and PD of extinct birds. (A)** Diagram of our classification of species groups [130,000 BP (All), 1500 CE (IUCN), current (Cur), and future (Fut)] and the different time

period comparisons used for assessing FD and PD loss. **(B)** Distribution of extinct bird species, separated into island endemics and continental species. In each case, the proportion of pre-1500 CE and post-1500 CE extinctions are