

# ECOGRAPHY

## Research article

### The what, how, and why of trait-based analyses in ecology

Thomas Guillerme<sup>1</sup>✉, Pedro Cardoso<sup>2,3</sup>, Maria Wagner Jørgensen<sup>4</sup>,  
Stefano Mammola<sup>3,5,6</sup> and Thomas J. Matthews<sup>4,7</sup>

EDITOR'S  
CHOICE

<sup>1</sup>School of Biosciences, The University of Sheffield, Sheffield, UK

<sup>2</sup>CE3C – Centre for Ecology, Evolution and Environmental Changes, CHANGE – Global Change and Sustainability Institute, Faculty of Sciences, University of Lisbon, Lisbon, Portugal

<sup>3</sup>Laboratory for Integrative Biodiversity Research (LIBRe), Finnish Museum of Natural History (Luomus), University of Helsinki, Helsinki, Finland

<sup>4</sup>School of Geography, Earth and Environmental Sciences and Birmingham Institute of Forest Research, University of Birmingham, Birmingham, UK

<sup>5</sup>Molecular Ecology Group (MEG), Water Research Institute, National Research Council (CNR-IRSA), Verbania Pallanza, Italy

<sup>6</sup>National Biodiversity Future Center, Palermo, Italy

<sup>7</sup>Centre for Ecology, Evolution and Environmental Changes/Azorean Biodiversity Group/CHANGE – Global Change and Sustainability Institute and Universidade dos Açores – Faculty of Agricultural Sciences and Environment, Angra do Heroísmo, Açores, Portugal

**Correspondence:** Thomas Guillerme ([t.guillerme@sheffield.ac.uk](mailto:t.guillerme@sheffield.ac.uk))

#### Ecography

2025: e07580

doi: [10.1111/ecog.07580](https://doi.org/10.1111/ecog.07580)

Subject Editor: Barry Brook

Editor-in-Chief: Miguel Araújo

Accepted 4 February 2025



Functional diversity is increasingly used alongside taxonomic diversity to describe populations and communities in ecology. Indeed, functional diversity metrics allow researchers to summarise complex occupancy patterns in space and/or time across communities and/or populations in response to various stressors. In other words, investigating what, how, and why something is changing in an ecosystem by looking at changes of patterns under a certain process through a specific mechanism. However, as the diversity of functional diversity metrics and methods increases, it is often not directly clear which metric is more readily appropriate for which question. We studied the ability of different functional diversity metrics to recover patterns and signals from different processes linked to common assembly mechanisms in community ecology, such as environmental filtering, competitive exclusion, equalising fitness, and facilitation. Using both simulated data and an empirical dataset affected by more complex and nuanced mechanisms, we tested the effectiveness of different space occupancy metrics to recover the simulated or empirical changes. We show that different metrics perform differently when trying to capture signals from different approximations of common mechanisms relative to no mechanism at all (null). For example, competition was harder to disentangle from the null mechanisms compared to facilitation in our simulations. This emphasises the importance of not using a one-size-fits-all metric. Instead, researchers should carefully consider and test whether a particular metric will be effective in capturing a pattern of interest.

**Keywords:** disparity, dissimilarity, functional diversity, mechanisms, patterns, processes



[www.ecography.org](http://www.ecography.org)

© 2025 The Author(s). Ecography published by John Wiley & Sons Ltd on behalf of Nordic Society Oikos

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.

## Introduction

In the last two decades, there has been a progressive expansion in ecology and evolution studies from taxonomic-oriented approaches, with species as the focal point, to functional approaches that place species-specific characteristics (traits) at the centre of analyses (Violle et al. 2007, Mammola et al. 2021, Palacio et al. 2022). Although numerous definitions exist (Dawson et al. 2021), we here consider a trait to be any observable characteristic (e.g. morphological, anatomical, ecological, physiological, behavioural, phenological) measured on individual organisms at any level, from genes to whole organisms. A trait-based approach has two advantages for answering core questions in ecology and evolution. First, it allows a deeper understanding of the mechanisms generating biodiversity patterns, by putting organisms' traits at the centre of natural selection rather than the organisms themselves. Second, it allows comparisons across subdisciplines in biology, while facilitating the conceptualisation of general principles broadly valid in space (e.g. unrelated species pools) and time (e.g. anatomical traits are comparable between palaeontology and ecology, which is not always the case with species; Luza et al. 2023). Already used routinely in palaeontology (Raup 1961, Gould 1991, Foote

1995, Guillerme et al. 2020a), this trait-focused ecology and evolution is unlocking the possibility to answer a broad range of questions in disciplines as diverse as community ecology (McGill et al. 2006), biogeography (Violle et al. 2014), conservation biology (Chichorro et al. 2022), micro- (Chapin et al. 1993) and macro-evolution (Guillerme et al. 2023), and applied fields (e.g. agronomy Martin and Isaac 2015). This is because trait-based approaches closely align with the general analytical framework proposed by Anand (1994) to answer three sequential questions: 'what?' (describing the pattern), 'how?' (describing the process) and 'why?' (understanding the mechanism) (Box 1).

Although some previous work has been done in understanding patterns in the context of mechanisms and processes (Novack-Gottshall 2016a, 2016b), we argue that often the pattern description follows some previously used framework without a proper evaluation of its adequacy to answer the how and why. The choice of the tool or metric used to describe the pattern is crucial for allowing us to understand the process and the mechanism. Using an inappropriate metric for describing a pattern can lead to biased conclusions. For example, let us imagine one is interested in understanding how two populations compete with each other for a resource (the mechanism – 'why?') on two islands, one home to both

### Box 1. What, how, and why: pattern, process, and mechanism.

Anand (1994) conceptualised a general analytical framework for answering scientific questions in ecology and evolution, based on three sequential questions: 'what?' (describing the pattern), 'how?' (describing the process), and 'why?' (understanding the mechanism). This framework can be effectively applied to trait-based analyses:

What? Pattern description corresponds to the steps needed to collect and summarise data to answer a question of interest (i.e. 'how?' and 'why?' as defined below). Usually this consists of: 1) collecting target traits at the focal level (e.g. genes, individual, population, species; Violle et al. 2007); 2) arranging these traits into some kind of trait space (Guillerme et al. 2020a, Mammola et al. 2021); and 3) using one or more unidimensional or multidimensional statistic to summarise properties of the trait space. For unidimensional spaces (i.e. distributions), these statistics can be as simple as the mean and the standard deviation (e.g. community weighted mean). For multidimensional spaces, the statistics are usually named disparity metrics or indices (Guillerme et al. 2020a) or functional diversity metrics (Mammola et al. 2021), but all attempt to capture some pattern of interest in the trait space.

How? Process description can be seen as linking the pattern of interest ('What?') to some dynamic element. This can be a punctual change. For example, change of the pattern under a certain condition, such as how traits X differ between two habitats (Martínez et al. 2021). But it can also be one or more continuous or ordinal changes, such as how the pattern X changes in space, time, and/or along ecological gradients (Belmaker and Jetz 2013, Lamanna et al. 2014, Bjorkman et al. 2018, Jarzyna and Jetz 2018, McLean et al. 2021). The distinction here might seem trivial but it is significant: usually, the process designates the change of the pattern, not the change of the traits. Although the traits are what is really changing, researchers will usually analyse some emergent property of the trait aggregation as described by the change in statistic values between two or more conditions.

Why? Mechanism description is then about linking the pattern and the process to some biological properties (sometimes referred to as 'rules') and is at the core of answering the biological question at hand. In most cases, this is what researchers are actually trying to understand. For example, one might be interested in understanding the effect of climate change on some trait (Boonman et al. 2022). In fact, most researchers work on understanding the causal link between variables (i.e. 'why' are 'what' and 'how' linked). This can be very useful, for example, to provide predictions about the past or the future. Ultimately, studying the mechanisms ('why') is often the reason why researchers get funded (or not).

Note that the distinction between what, how, and why is not categorical in its nature and is often nuanced, with patterns and process or process and mechanisms sometimes being used interchangeably to describe the same questions.

populations and another one home to only one of the population (the process – ‘how?’). In such a scenario, measuring the occupied area (i.e. km<sup>2</sup>) of each population on each island (a pattern – ‘what?’) will not be the most appropriate way to understand the potential competition between these populations because it might not give information on whether the two populations actually interact in some way – a very likely condition for competition to take place. In this simplistic example, the functional overlap between populations (‘what and how’) might be more appropriate. See for example [Carvalho and Cardoso \(2020\)](#) for a discussion of how Darwin’s finches share resources depending on the existence or not of competition.

Through a simulation exercise, we analysed different patterns (what) across different processes (how) approximating different mechanisms (why) of interaction between organisms: equalisation, filtering, facilitation, and competition ([Fig. 1](#), [Table 2](#)). Our goal is to test the relative performance of different metrics to capture the patterns of different mechanisms by comparing the scores of various metrics under specifically approximated mechanisms and the absence of any specific mechanism (null mechanism). We expect that our ability to capture a pattern does not only depend on the choice of metric (what) but also on the process and mechanism at hand (how and why). To apply our framework to a realistic context, we also used an empirical dataset of Hawaiian bird traits and looked at how anthropogenic pressures have shaped the trait space (mechanism; why) based on extinction events on Hawaiian islands before and after the year 1500 (process; how), and how different metrics (pattern; what) can lead to different interpretations of the data. We show that, with a fixed process and mechanism approximation, the choice of the statistic to describe the pattern (the disparity or functional diversity metrics) has a great impact on the interpretation of the data, with the exact same data sometimes being

interpreted in opposite ways leading to sometimes opposite potential interpretations.

## Material and methods

### Simulating trait space patterns

First, we simulated multiple independent random time dependent traits (Brownian motion) under a model where lineages only speciate (no extinction; i.e. a pure birth speciation model) until reaching 200 observations in R ([R Core Team 2024](#)) using `treats` ([Guillerme 2024](#)). We simulated either 2, 4, or 8 independent (uncorrelated) Brownian motion traits (equivalent to 2, 4, or 8 dimensional traits). Note here that we refer to traits as any measurable aspect of an organism (sensu [McGill et al. \(2006\)](#) – but see [Dawson et al. \(2021\)](#) for nuances in what researchers perceive as traits). These aspects are often expressed in one dimension but can in fact be described in any number of dimensions. For example, for a trait defined as ‘leaf insertion angle’, this can be measured in one dimension (an angle in degrees) or three dimensions (the same angle expressed as the trigonometric relation of three sides of a triangle). The observations on which the traits are measured can represent any biological focal entities, e.g. tips in a phylogenetic tree, individuals, species, populations, and OTUs ([Table 1](#)). This simulation approach resulted in a neutral null model of trait evolution with no effect of competition, extinction, selection, or other processes (sensu [Bausman 2018](#)). We refer to this as the ‘non-stressed trait space’.

### Applying stressors to the trait space

We then applied five different stressors to the trait space with different intensities. The five different stressors are effectively

Table 1. Glossary and equivalence of terms used in this paper; adapted from [Guillerme \(2018\)](#) and [Mammola et al. \(2021\)](#). \*Note that our definition of a trait is a generalisation of [McGill et al. \(2006\)](#), [Violle et al. \(2007\)](#), and [Dawson et al. \(2021\)](#)’s suggestions expanded to include traits beyond the ones common in functional ecology (e.g. also in macroevolution, palaeontology, and microbiology).

Term in this paper	Definition	In ecology	In macroevolution
Trait space	Matrix ( $n \times d$ ) with a structural relation between rows and columns	Functional space, morphospace, etc.	Morphospace, trait space, etc.
Observations	Rows ( $n$ )	Taxa, field sites, environments, etc.	Taxa, specimen, populations, elements, OTUs, etc.
Dimensions	Columns ( $d$ )	Traits, ordination scores, distances, etc.	Traits, ordination scores, distances, etc.
Traits	Columns subset ( $b \times n$ ; $b \leq d$ )	Uni/multidimensional characteristic*; e.g. body mass, geographical location	Uni/multidimensional characteristic; e.g. landmark coordinates
Group	Rows subset ( $m \times d$ ; $m \leq n$ )	Treatments, phylogenetic group (clade), etc.	Clades, geological stratum, etc.
Metric	Statistic (i.e. a measure)	Dissimilarity index or metric, hypervolume, functional diversity, etc.	Disparity metric or index
Stressors	An algorithm removing a proportion of the observations	Introduction of invasive species, change of landscape use, climate change, pollution, experimental design, etc.	Mass extinction, tectonic change, climate change, etc.
Stressors’ intensity	The amount of observations removed (percentages)	Proportion of species affected, proportion of temperature change, etc.	Severity of mass extinction, number of co-occurring species, etc.

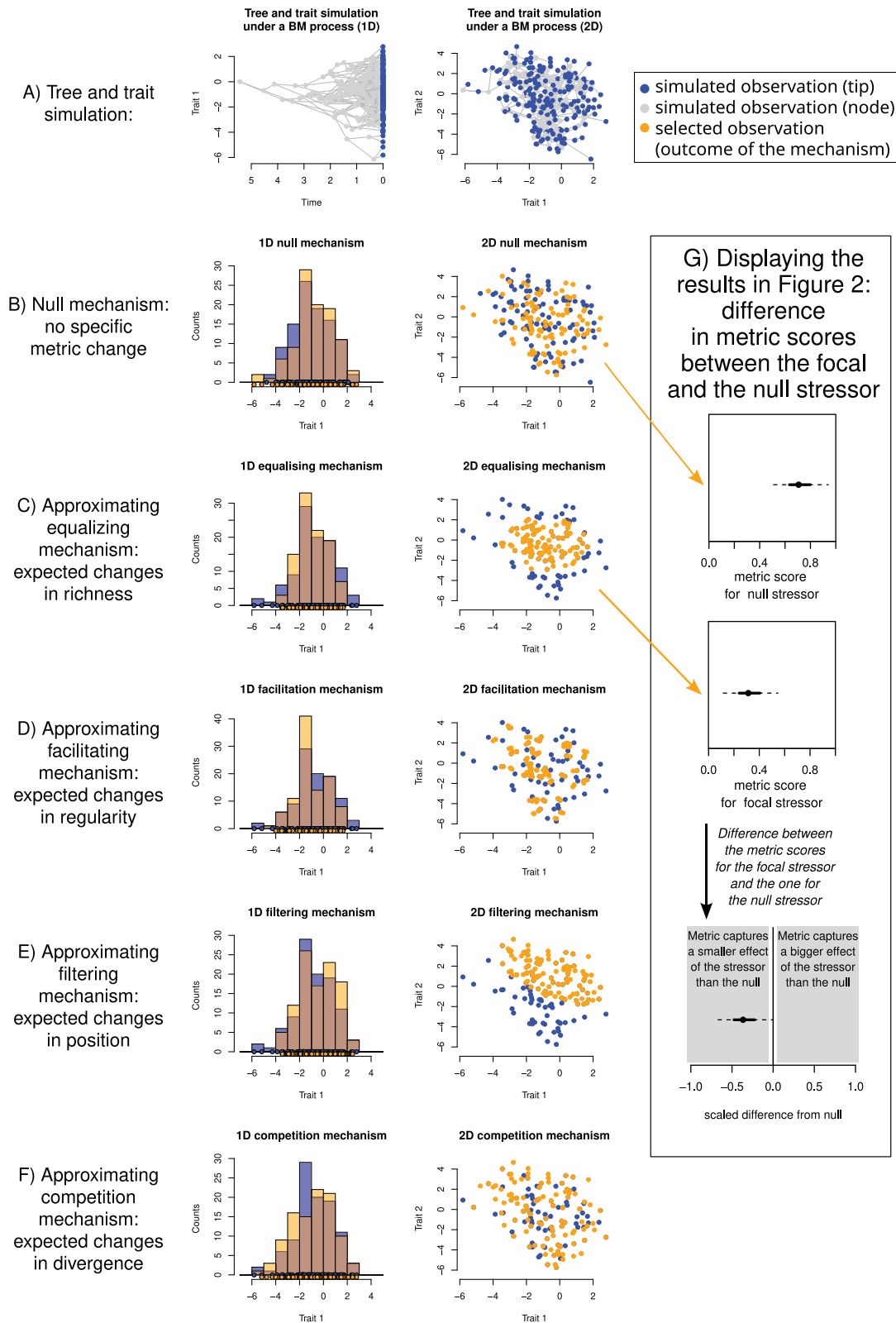


Figure 1. Illustration of the simulation protocol used in this paper. (A) We simulated a pure birth-tree with a trait evolving under a Brownian motion process until reaching 200 observations resulting in a trait space of 2, 4, or 8 dimensions. The first plot on row (A) illustrates one such simulation in one dimension through time (with time on the horizontal axis and trait value for one dimension on the vertical axis). The second plot on the same row shows the same results but with a two dimensional trait represented here. We then applied different stressors to the resulting trait space. We represent the elements removed by the stressors in blue and the ones kept in the trait space in orange.

Figure 1. Continued.

The first column illustrates the distribution of the elements in one dimension with the overlap between the removed (blue) and kept (orange) elements in brown (note we did not measure any metric on 1D spaces). The second column illustrates the same distributions in 2 dimensions. We do not illustrate the distributions in 4 or 8 dimensions here. We used the following stressors: (B) null mechanism: randomly removing a proportion of the observations; (C) approximating equalising mechanism: removing observations on the edge of the distribution; (D) approximating facilitation mechanism: removing observations to reduce the distance between pairs of observations (increasing local density; decreasing evenness); (E) approximating filtering mechanism: removing observations from one extreme of the distribution; (F) approximating competition mechanism: proportionally removing observations from the centre of the distribution. The expected changes indicate the main type of metric affected by the stressor (following the aspects in Mammola et al. 2021). (G) we presented the results of the simulations as the scaled metric scores for the approximated mechanisms (C–F) relative to the null mechanism. The dot represents the median metric score and the thick and dashed lines the 50 and 95% confidence intervals, respectively.

five specific algorithms removing either 20, 40, 60, or 80% of the data (the intensities of the stressors, resulting in trait spaces with, respectively, 160, 120, 80, and 40 observations). We used the following specific algorithms (Fig. 1, Table 2; all algorithms, except ‘increasing evenness’, were previously described in Guillerme et al. 2020b):

- Random removal: by randomly removing 20, 40, 60, or 80% of the data. This approximates our null mechanism.
- Decreasing size: by removing the required amount (20, 40, 60, or 80%) of data away from a distance (radius)

This is used to establish a reference to be compared to the other mechanisms and to test how removing elements in a specific way influences the metric scores compared to removing them randomly. We chose this algorithm to approximate the absence of any specific stressor (i.e. the removal of data only changes the number of observations, not the properties of the trait space in a systematic way).

Table 2. Description of the stressors applied to the simulated data and the mechanism each approximated. Here we distinguish between the biological mechanism we are trying to simulate and the algorithmic one we used to simulate it. We also provide a more detailed description of the stressor. Note that, in biological data, we don't expect any of the mechanisms to act on communities alone (e.g. equalising fitness and facilitation can both act on the distribution of species traits). Nor do we expect their effects to be unidirectional (for example, equalising fitness can happen both by removing the edges or changing the position of a trait distribution), thus, the type of metric expected to recover the mechanism can be variable. These mechanisms serve as a simplified description of reality for the narrative purpose of this paper. \* Richness, divergence, and regularity are dimensions of functional diversity metrics that have been proposed to reflect the sum of differences, the average differences, and how regular the differences are among observations, respectively (Pavoine and Bonsall 2011, Tucker et al. 2017, Mammola et al. 2021).

Approximated mechanism (biological)	Used mechanism (algorithm in dispRity::reduce.space)	Stressor description	Type of metric expected to recover the mechanism
Null mechanism	Random removal ('random')	No overall systematic change in observations' traits or community structure except for the reduction in the number of observations	Metrics sensitive to the amount of data
Equalising fitness	Size change ('size')	When observations with trait combinations that are more extreme (i.e. observations located away from the centre of the trait space) are disadvantaged due to some resource concentration gradient (Chesson 2000, Barot 2004)	Metrics mainly capturing changes in richness* (but less in divergence* and not in regularity*)
Facilitation	Density change ('density')	When observations with similar trait combinations (i.e. located close together in the trait space) are more likely to survive the stressor (Bruno et al. 2003, Danet et al. 2024). In other words, this simulates the idea that some observations are more likely to be resilient if they are present in a community of observations (e.g. a community of species) with shared traits rather than the opposite	Metrics capturing changes in regularity* (but not richness* and divergence*)
Environmental filtering	Position change ('position')	When an increase in observations' trait similarity happens through some strong abiotic constraints (Cornwell et al. 2006). For example, some trait combinations become more and more unlikely due to some environmental constraints (i.e. some regions of the trait space become unsuitable)	Metrics capturing changes in position (and richness*)
Competitive exclusion	Evenness change ('evenness')	When functionally similar observations compete more intensively with one another than with functionally dissimilar observations ('competitive exclusion principle'; Hardin 1960). At the extreme, it implies that only dissimilar observations will coexist ('limiting similarity principle'; MacArthur and Levins 1967)	Metrics capturing changes in divergence* (but not richness* and regularity*)

$\rho$  of the centre of the trait space.  $\rho$  is estimated for each trait space to match the required intensity. For example, for one trait space, the algorithm estimated  $\rho_{20}$  to be the minimum radius excluding 20% of the observations,  $\rho_{40}$  the one removing 40%, etc. (with  $\rho_{20} > \rho_{40} > \rho_{60} > \rho_{80}$ ). This approximates our equalising mechanism. We chose this algorithm to approximate equalisation assuming that a stressor could increase the probability of extinction for observations with more rare trait combinations (i.e. observations on the edges of the trait space).

- Increasing density: by removing the required amount of pairs of points with a variable pairwise distance of at least  $D$ .  $D$  is a distance estimated for each trait space to correspond to the minimal Euclidean distance that encompasses at least  $n$  pairs of points corresponding to the intensity of the stressor. For example  $D_{20}$  is the distance that excludes any  $n_{20}$  pairs of points that are a distance of at least  $D_{20}$  from each other,  $D_{40}$  is the distance that excludes  $n_{40}$  pairs, etc. (with  $D_{20} > D_{40} > D_{60} > D_{80}$ ). Note that this algorithm is not directly based on the change in average density but rather on the change in pairwise distance between observations. This approximates our facilitation mechanism (i.e. the points left are only ones that are close to at least another point in space). We chose this algorithm to approximate the facilitation mechanism where a stressor could increase the probability of extinction for observations that are more isolated in the trait space. In other words, pairs of observations that share similar trait combinations are more common in the trait space than observations with dissimilar trait combinations (i.e. observations adjacent to each other in the trait space are more common).
- Shifting space: by removing the required amount of data from a distance (radius)  $\rho'$  of the observation with the maximum numerical value on all dimensions. This is similar to the decreasing size algorithm but, instead of choosing  $\rho$  to be the radius from the centre of the trait space,  $\rho'$  is a radius from the 'top right corner' of the trait space. In other words the centre of the radius  $\rho'$  is the observation with the highest numerical value on all dimensions, in a 2D representation this is the observation in the top right corner (with  $\rho_{20} > \rho_{40} > \rho_{60} > \rho_{80}$ ). This approximates our filtering mechanism. We chose this algorithm to approximate a filtering mechanism assuming a stressor that could increase the probability of extinction for observations further from some trait combination optimum. This is similar to our equalising mechanism approximation but with the optimum not being in the centre of the trait space (a region with a high density of observations) but with the optimum being in a corner of the trait space (a region with low density).
- Increasing evenness: by resampling the proportion of data (i.e. 20, 40, 60, or 80%) but with skewed resampling probabilities. In brief, this algorithm reduces the probability of resampling observations in regions of the trait space that have many observations and increases it in regions that have few observations. This is done by

selecting  $B$  discrete categories to summarise the distribution – i.e. bandwidths – using Silverman's 'rule of thumb' (*bw.nrd0* function in R; Silverman 1986). Each discrete category  $B$  has an observed probability of sampling of  $b$  (the proportion of observations in the category) and each observation then gets a probability of resampling scaling with the intensity of the stressor (the proportion of what to remove  $i$ ) of  $i \times (1 - b)^p$ . Where  $p$  is a factor increasing the scaling power of the algorithm (here we arbitrarily used  $p = 3$  which was a relatively low scaling value that still resulted in visible changes observable in 2D). In other words, observations in categories with a low density ( $b < 0.5$ ) were more likely to be resampled and the ones in categories with high density ( $b > 0.5$ ) were less likely to be resampled. This approximates our competition mechanism, where observations in dense regions of the trait space are more likely to be removed than in sparse regions of the trait space (or, inversely, observations are more likely to be 'preserved'/'surviving' in less dense regions). We chose this algorithm to approximate a stressor that could increase the probability of extinction for observations that share trait combinations. This could approximate the fact that observations that have similar trait combinations are more likely to go extinct due to competition (e.g. using the same resources).

For the decreasing size, increasing density, and shifting space stressors, we estimated the parameters  $\rho$ ,  $D$ , and  $\rho'$  recursively with the `disPRity` package to obtain the required amount of data to be removed (`disPRity::reduce.space`, Guillerme 2018, Guillerme et al. 2020b). Note that the algorithm used here to simulate the stressors depends, to some extent, on the distribution of the data for each trait space. The algorithms used can lead to similar stressor effects depending on the simulated data distribution. For example, 'increasing density' and 'decreasing size' stressors will lead to just what their names suggest if the observations are uniformly distributed (just an average increase in density for 'increasing density' and just an overall decrease in size for the 'decreasing size' stressors) but will have more or less similar effects (albeit not identical) in the case of a normal distribution (average increase in density and overall decrease in size for both 'increasing density' and 'decreasing size' stressors). These steps resulted in 3420 simulated trait spaces (171 simulated spaces  $\times$  5 stressors  $\times$  4 intensities of data removal). We used 171 replicates because that was empirically the smallest number of replicates required to reach a variance between replicates lower than 1% across all metrics (i.e. any additional replicates beyond 171 added less than 1% extra variance). See Table 2 for a biological description of these stressors and Fig. 1 for a visual description of them.

Note that in empirical data, depending on the distribution of the data, some mechanisms can lead to similar or dissimilar patterns. For example, if the data are normally distributed, the equalising mechanisms, by removing data on the edges of the distribution, also increases the density of the trait space (because normally distributed data are denser in the centre of

the distribution), in a similar way to the facilitation mechanism. However, if the data are distributed uniformly, this does not happen. Furthermore, in real-world scenarios, we do not expect these mechanisms to act in isolation of each other: multiple mechanisms may stress the observed data simultaneously, with cumulative or synergistic effects. However, this is not tackled here for both simplicity and to understand how the stressors work in isolation. Therefore the prediction of how each different mechanism approximation would affect the metrics (Table 2) should be taken as a cautionary guideline when dealing with empirical data where the distribution of the data might not be normal and where it could be affected by multiple mechanisms.

We repeated each simulation pipeline for 4 dimensional traits (generate a trait space and apply the stressor – Fig. 2), and for 2 and 8 dimensional traits (Supporting information). In each pipeline, the dimensions were simulated as uncorrelated. We limited our simulations to a relatively small number of dimensions due to the constraints of some of the metrics used (e.g. TPD::TPDsMean is only implemented for up to 4 dimensions, Carmona et al. 2019) but also to avoid dealing with the curse of dimensionality (Bellman et al. 1957). This curse changes the properties of space occupancy in a non-linear way depending on each observation's distribution and the number of dimensions. For example, the volume of a trait space (or hypervolume when using more than 3 dimensions) typically tends to zero in a high number of dimensions. However, the rate at which it approaches zero is not linear and depends on the distribution of each observation on every dimension. This makes it practically difficult to compare two randomly generated spaces with similar characteristics (e.g. for two spaces with 200 observations and 10 dimensions generated in the exact same way, one might have a hypervolume of nearly 0 and the other one of  $10^3$ ). Note also that, in our simulations, all dimensions have the same properties (i.e. same variance and distribution). This is often not the case in empirical cases (e.g. see our empirical example) where the dimensions have a decreased variance due to ordination techniques.

### Metrics for measuring trait space occupancy (aka functional diversity, dissimilarity, disparity)

We structured our simulations based on three aspects of diversity commonly captured by functional diversity metrics (trait space occupancy metrics; Mammola et al. 2021) using 12 metrics (Table 3): 1) richness, encompassing metrics reflecting the sum of differences among observations; (four metrics, equivalent to size metrics in Guillerme et al. 2020b); 2) divergence, encompassing metrics reflecting the average differences among observations (four metrics); and 3) regularity, encompassing metrics reflecting how regular the differences among observations are (four metrics, equivalent to density metrics in Guillerme et al. 2020b). Note that we focused here on three packages fully devoted to functional diversity analyses (FD Laliberté and Legendre 2010, BAT Cardoso et al. 2015, and TPD Carmona et al. 2019) in the R statistical environment (Mammola et al. 2021).

### Comparing and scaling the results to the random removals (null stressor)

To understand the ability of each metric to capture a mechanism of interest, we compared their scores relative to the same metric scores of a null mechanism. To ease interpretation of the results across a range of metrics with different orders of magnitudes of outputs and different random simulations, we first calculated the difference between the stressor of interest and the null stressor (e.g.  $\text{metric}_{\text{stressor } 20\%} - \text{metric}_{\text{null } 20\%}$ ). We then scaled these differences to be relative to the highest relative differences between all the simulations for that one metric and stressor (across the four levels of removal –20, 40%, etc.) This resulted in each metric being scaled between –1 and 1 where a negative difference can be interpreted as the random removal of elements leading to a higher metric score; a positive difference as the focal removal of elements (i.e. approximating equalisation, facilitation, filtering, and competition) leading to a higher metric score; and a difference near 0 meaning that both the null stressor and the focal stressor have a similar effect on the metric score (Fig. 1–3). We interpret this last scenario as the metric score being unable to clearly distinguish between a random removal of elements and a non-random one. To assess these poor performances further, we compared the distributions of the 171 scores for each metric with each reduction level for each non-null stressor to the null stressor using pairwise t-tests. The results of these tests are displayed in Fig. 2 (for the non-null stressor, grey distributions are not clearly distinguishable from the null stressor).

### Measuring the effect of the strength of the stressor

Finally, we measured the effect of the different strengths of the stressors (i.e. removing 20, 40, 60, and 80% of the data) by fitting a linear regression using the scaled metric difference to the null score as a function of the amount of data removed. For each of these models, we reported the variance explained by the model (adjusted  $R^2$ ) in the Supporting information and whether the slope was clearly distinguishable from 0 (Fig. 2; sensu Dushoff et al. 2019). This allowed us to see if there was any predictable linearity in the metric scores. In other words, testing whether the metric scores relative to null removals were linearly scaling with the amount of data removed.

### Empirical data

We applied all metrics to an empirical dataset of bird extinctions on six Hawaiian islands (Hawaii, Kauai, Lanai, Maui, Molokai, Oahu), an archipelago which has suffered large numbers of anthropogenic extinctions due to a range of extinction drivers (Walther and Hume 2022). A prehistoric species list (avifauna known to be present prior to human colonisation of the islands), historic species list (avifauna known to be present at 1500 CE), and extant species list of the native community for each island were taken from Matthews et al. (2023) along with trait values for the extinct species (*Asio flammeus* was removed from the historic and

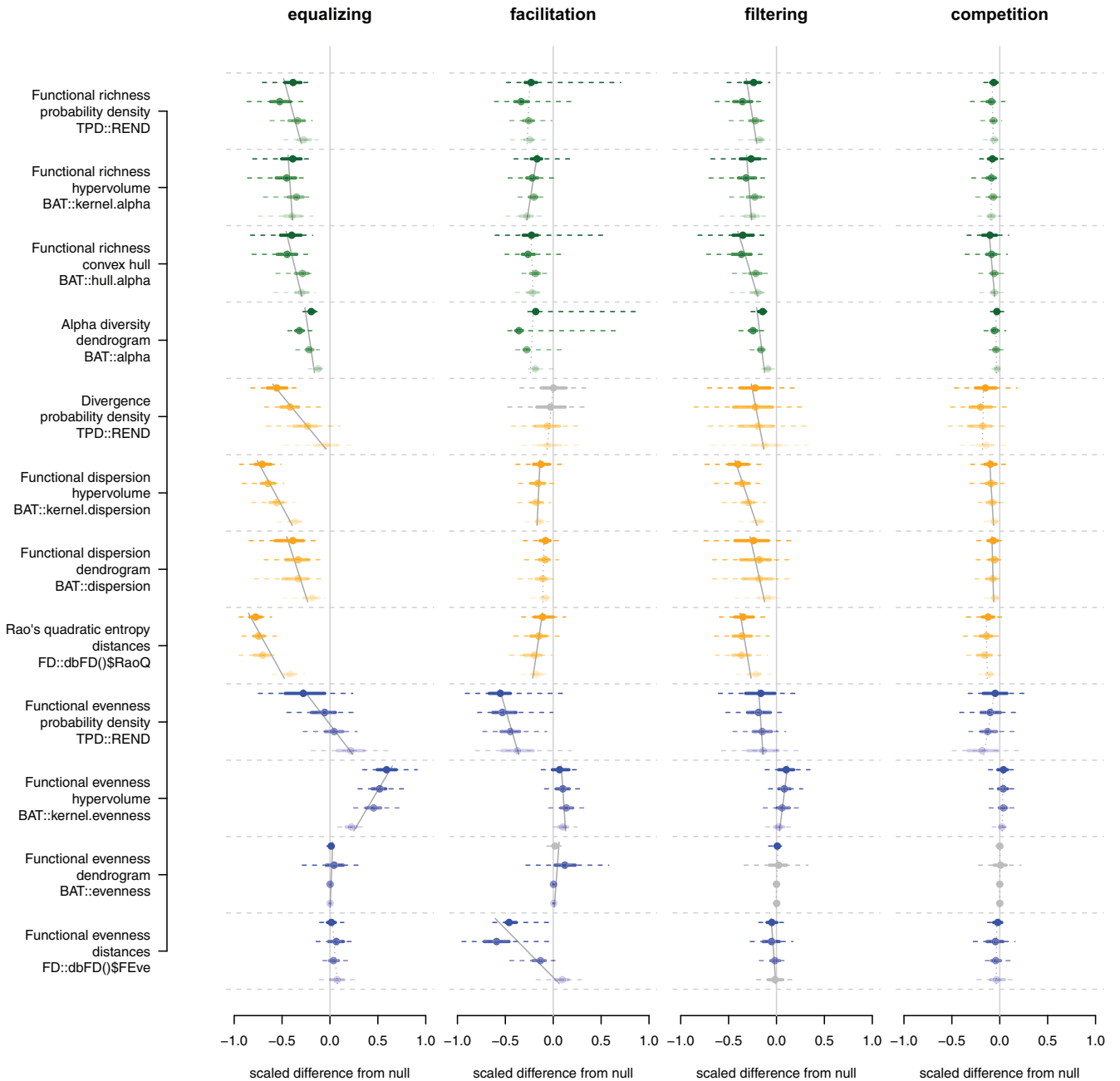


Figure 2. Simulation results: the y-axes represent the different metrics tested (sorted by categories). The different columns represent the different stressors. The x-axes represent the metric values centred on the random changes and scaled by the maximum value for each metric between the four stressors. Negative and positive values signify that the metric score for the stressor of interest is lower/higher than the one from the random stressor. The dots represent the median metric value, the full line their 50% confidence interval (CI), and the dashed line their 95% CI. The colours are used here to visually separate the metrics by categories (green = richness, yellow = divergence, blue = regularity); the colour gradient within each row corresponds to a removal of respectively 80, 60, 40, and 20% of the data (from top to bottom). The grey dots and corresponding CI lines represent distributions of metric scores not clearly distinguishable from the random metric scores (paired t-test p value > 0.05). Grey lines in the background across the distributions of different removal amounts represent the following fitted linear model 'centred and scaled metric score function of the amount of data removed'. Dashed thin grey lines represent non-significant linear models (slope's p value > 0.05). The model results (including R<sup>2</sup> and t-statistics) are available in the Supporting information. Similar figures are available in the Supporting information for 2 and 8 dimensions.

Table 3. Functional diversity metrics tested in our simulations. \* The TPD package currently only allows calculations for trait spaces of up to four dimensions. This metric was excluded from trait spaces with more than four dimensions.

Aspect	Method	Name	Description	R function	Reference
Richness	Trait probability density	Functional richness	Sum of cells where trait probability density > 0	TPD::REND*	Carmona et al. 2019
	Kernel density hypervolume	Functional richness	Size (volume) of the hypervolume	BAT::kernel.alpha	Mammola and Cardoso 2020
	Convex hull	Functional richness	Volume of the convex hull	BAT::hull.alpha	Cornwell et al. 2006
	Dendrogram	Alpha diversity	Total branch length of the functional dendrogram	BAT::alpha	Petchey and Gaston 2002, Cardoso et al. 2015
Divergence	Trait probability density	Divergence	Overlap between the TPD function of a single observation and the TPD function of the whole set of observations	TPD::REND*	Carmona et al. 2019
	Kernel density hypervolume	Functional dispersion	Average distance between an observation and a sample of random points within the probabilistic hypervolume	BAT::kernel.dispersion	Mammola and Cardoso 2020
	Dendrogram	Functional dispersion	Average dissimilarity between any two observations in the dendrogram	BAT::dispersion	Cardoso et al. 2015
	Distances	Rao's quadratic entropy	The mean Euclidean distance between every two random pairs of species	FD::dbFD()\$RaoQ	Botta-Dukát 2005
Regularity	Trait probability density	Functional evenness	An index of the spread of the trait distribution compared to an even spread of the distribution (e.g. the height of bins in a histogram)	TPD::REND*	Carmona et al. 2019
	Kernel density hypervolume	Functional evenness	Overlap between a hypervolume and an imaginary hypervolume where dimensions are evenly distributed within their possible range	BAT::kernel.evenness	Mammola and Cardoso 2020
	Dendrogram	Functional evenness	Regularity of abundances and distances between observations in the tree.	BAT::evenness	Cardoso et al. 2015, 2024
	Dendrogram	Functional evenness	Regularity of abundances and distances along the minimum spanning tree linking all observations	FD::dbFD()\$FEve	Villéger et al. 2008, Laliberté and Legendre 2010

extant species lists as its colonisation status is uncertain). We extracted traits for extant species from the AVONET database (Tobias et al. 2022). We used nine morphological traits associated with dietary/foraging preference and dispersal ability (Pigot et al. 2020, Sheard et al. 2020) to represent the functional diversity of the communities: mass; beak length (culmen); beak length (nares); beak width; beak depth; tarsus length; wing length; the length from the first secondary feather to the tip of the longest primary feather (Kipp's distance); and tail length. Traits were log transformed and standardised to a mean of 0 and standard deviation of 1, prior to analyses.

The full trait dataset comprised 118 bird species, all native species known to have existed on these islands over the last 125 000 years. It is known that 55 species have gone extinct prior to 1500 CE and 26 after 1500 CE. Note that we used the extinct species data as presented in Matthews et al. (2023), and thus we do not include Hawaiian species declared extinct by the IUCN in 2024. For our analyses, we focused on two time periods: 1) the avifauna present in 1500 CE (the 'historic dataset') and 2) the current native avifauna ('extant dataset'). We represented the distinct reductions in species richness preceding these two time periods as our two stressors, the first one representing all pre-1500 CE extinctions

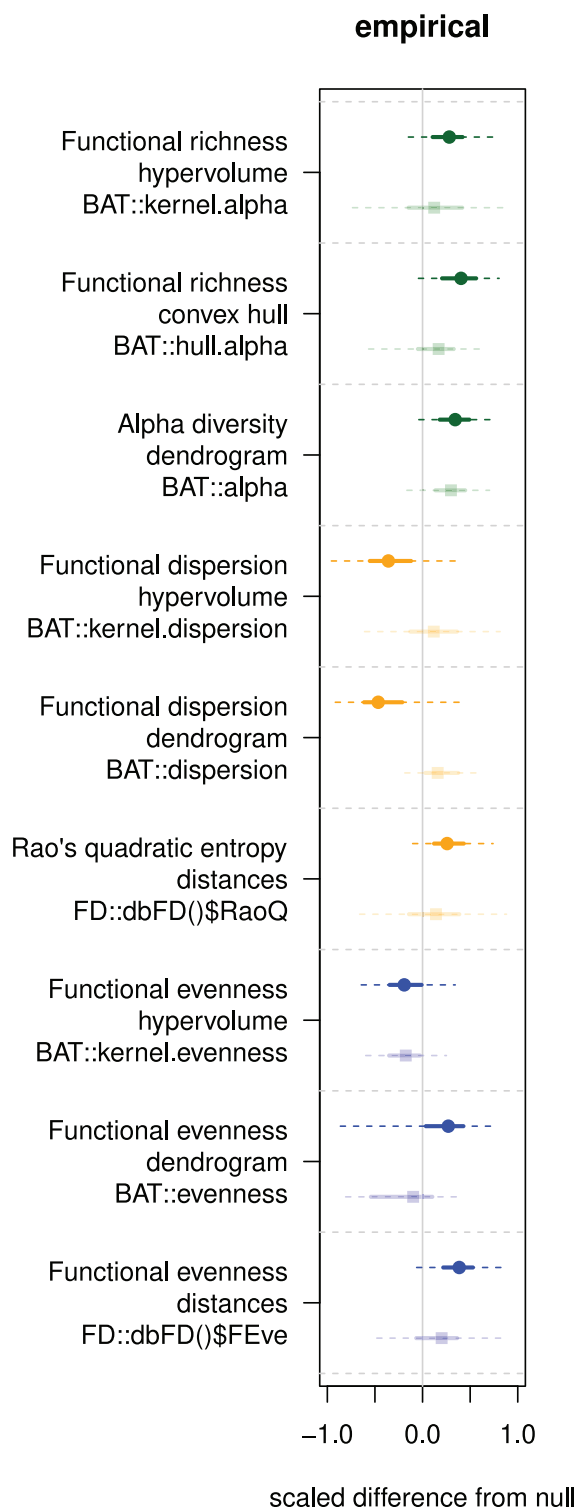


Figure 3. Empirical results: the metrics and the scaled changes are measured in the same way as in Fig. 2 (green = richness, yellow = divergence, blue = regularity). Plain circles represent the metrics for the species present on Hawaii in 1500 (historic data; 46% removal – 63 species present) and faint squares for the species currently present on Hawaii (extant data; 69% removal – 37 species present). The model results (including  $R^2$  and  $t$ -statistics) are available in the Supporting information.

(i.e. 118–55 species; a species richness reduction of 46%) and the second all the extinctions that have occurred until the present (i.e. 118 – (55 + 26) species; a species richness reduction of 69%).

To build the community trait space, we undertook a PCA including all 118 species, which we then subset to calculate the trait space for the historic and extant datasets. We selected the first five axes to represent the trait spaces that explained at least 95% of the variance in each specific group (historic and extant species – Supporting information; following the recommendations of Pigot et al. 2020 and the dimensionality selection method from Guillerme et al. 2023). We applied the same procedure as for the simulated data by simulating a null mechanism to compare to the observed ones by randomly removing 55 and 81 species for each stressor and scaling the results proportionally to this null mechanism (as described above).

## Results

The ability of different metrics to capture the different patterns (and thus approximate the mechanisms) was highly variable. It ranged from metrics capturing no clear pattern (e.g. functional evenness based on the dendrogram method for competition) to metrics clearly capturing one specific pattern (e.g. divergence based on probability density for the equalising mechanism). Although most metrics captured a decrease in metric score relative to the amount of removed data, some metrics resulted in a score increase (e.g. the functional evenness based on a hypervolume for the equalising mechanism; positive values in Fig. 2) or non-linear responses (e.g. alpha diversity based on a dendrogram for the equalising mechanism represent distributions of metric scores; Supporting information). Furthermore, although we predicted some more common changes in aspects of functional diversity for some specific mechanisms, our simulations show that most metrics under most aspects (richness, divergence, or regularity) capture changes of trait space occupancy under any mechanism.

For the empirical data, we again see a range of different possible interpretations depending on the metric describing the pattern (Fig. 3, Supporting information). For example, all richness metrics (green; Fig. 3) pick up a similar relative increase (i.e. relative to the null stressor) in richness for both the historic and extant datasets. However, the results are occasionally contrasting between the historic and extant datasets, with a respective decrease and increase with some divergence metrics (e.g. functional dispersion based on dendrograms or hypervolumes). Finally, some metrics pick up a decrease for both datasets (e.g. functional evenness based on hypervolumes – although the 95% confidence intervals overlapped with 0). Note that this apparent global increase in most metric values may appear counterintuitive given that 46 and 69% of species went extinct over the two time periods. This is due to the scaling of the metrics compared to random equivalent extinctions. That is, Fig. 3 is not displaying absolute changes in trait space occupancy, but rather changes in

trait space occupancy relative to a crude neutral model where all species are equally likely to go extinct.

## Discussion

### Capturing approximated mechanisms with different metrics

We tested 12 trait space occupancy metrics on simulated and empirical datasets to assess how each metric captures patterns of trait space changes based on stressors approximating ecological and evolutionary mechanisms. Our results show that different metrics capture different patterns (what) leading to inferring different processes (how) that can be proposed to explain different mechanisms (why). Therefore, the choice of trait space occupancy metric is essential to accurately describe a pattern of interest. We found that different metrics can capture changes in pattern relative to the null mechanisms in different ways depending on the approximated mechanism of interest. For example, for the approximated equalising mechanism, most metrics capture a pattern clearly smaller than the one from the null mechanism (i.e. randomly removing elements has a greater effect on the metric than removing elements following the mechanism). However, some metrics (e.g. functional evenness based on the hypervolume) capture a pattern clearly greater than the one from the null mechanism. This means that changes in metric values can be counterintuitive and should always be interpreted in the context of the metric, data, and question at hand. Note also that, as mentioned above, different dimensionalities can have different effects on the metrics, depending on both the trait space and the metric nature (Bellman et al. 1957).

For example, say one is interested in looking at changes in functional diversity using a divergence metric calculated with the probability density method (e.g. TPD::REND). They can apply this metric to their trait space and get two relative metric scores of 0.7 and 0.65 (pattern) for two levels of extinction (process). The interpretation of this score as a change in functional diversity will depend on the worker's assumption of which mechanism affected the trait space. In our simulations, if the mechanism approximated was the equalising one, the pattern can clearly be interpreted as a change in functional diversity, but if the mechanism is a facilitation one, the pattern can be meaningless (i.e. not clearly distinguishable from a null mechanism – Fig. 2 shows this example by contrasting columns 1 and 2 in row 5). In this made-up example, the procedure is done in reverse order where the pattern is specified first, and then the process and the mechanism ('what process and mechanism are needed to interpret pattern X?'). In Box 2 we suggest an approach that specifies the process and the mechanism first ('how do I interpret pattern X under a specific process and mechanism?').

Regarding the different approximated mechanisms, it is interesting to note that our different approximations lead to different general behaviours of the metrics: for example the range of variance in the metric values is greater for the

equalising mechanism and smaller for the competition mechanism (albeit nearly always distinguishable from the null mechanism). These differences in metric variance across mechanisms suggest that some mechanisms lead to more subtle changes or/and that our approximation of these mechanisms can be variable.

### Empirical results

The empirical results are also intriguing in that the loss of Hawaiian birds' functional diversity due to extinction is not greater than expected by chance (at least for certain metrics; Fig. 3), because one would expect that extinctions target specific areas of functional space (i.e. birds with specific trait combinations). This contrasts with studies that have shown that species with certain traits (e.g. larger species) are more likely to have gone extinct and thus that functional diversity loss is larger than through random extinctions (Sayol et al. 2021, Matthews et al. 2022). This illustrates the importance of positing the mechanism and process of interest as well as the pattern to capture it before analysing the results. For example, for the divergence metrics using hypervolume or dendrogram methods, we recovered changes in trait space that are smaller than expected by chance when focusing on all extinctions and larger than expected by chance when considering only pre-1500 extinctions. This could lead to the interpretation that the trait space was not changed under any specific mechanism up to 1500 (i.e. null mechanism) but that, since then, specific mechanisms have been reducing specific areas of the trait space (e.g. filtering mechanism). However, when using a different metric (say the Rao's quadratic entropy based on distance matrices), the results would indicate that a reduction of specific areas of the trait space could have also occurred before 1500, leading to very different interpretation of the results.

Our results show that the loss of functional diversity for several metrics was not more than expected by chance. This could be due to the taxonomic distribution of the extinct birds: these Hawaiian data comprise approximately 58% passerines while the majority of anthropogenic extinctions have involved non-passerines (approximately 75%). Non-passerines tend to be larger and possess more unique morphological traits (thus resulting in relatively large reductions in functional diversity). However, in Hawaii, where the same proportion of known species were passerines (approx. 58%), the opposite is true: the majority of extinctions (approx. 60%) have been of passerine species. It is important to note that we are not implying that there existed little morphological variation within extinct Hawaiian passerines – indeed, many of these extinctions involved the honeycreepers, a famous adaptive radiation involving substantial evolution of beak morphology (Walther and Hume 2022) – but simply that this morphological variation is less than that observed across all birds. Given that the proportion of known Hawaiian species that are/were passerines roughly matches the proportion of extinct species that were passerines, our results indicate that these extinct species were more functionally similar than expected, thus relatively increasing trait diversity (i.e. relative

## Box 2. So which metric do / choose?

TL;DR: it depends!

Here is a proposed step by step protocol that could help one decide what to measure by clearly specifying their mechanism, process and pattern of interest prior to any functional diversity or disparity measurements.

We have data of bird traits for native species in Hawaii through time (e.g. the birds known to be present in 1500 CE and those still extant today). We want to know how anthropogenic stressors (e.g. hunting, invasive species, land use change) have affected the trait diversity of native Hawaiian birds (Fig. 4).

### Frame the question

- Why? Think of the broad biological mechanism of interest. In our example, this could be asking why (due to what mechanism) has the community changed: the trait diversity has changed due to a filtering mechanism where anthropogenic stressors affect specific species with specific traits (e.g. through hunting, predation on eggs, by mammals).
- How? With the data we have at hand, identify the process we could analyse. In our example, this could be native bird species traits through time: how did filtering affect the native species after 1500?

Although asking why and how is a relatively standard step in the literature, we suggest that proposing them in this framework more easily leads to the next, often less asked question of what pattern to capture:

- What? Choose the appropriate pattern that would capture changes through time in relation to a filtering mechanism. In our example, this could be a continuous metric that measures either changes in the trait space position or/and changes in the trait space size. The idea being that filtering would remove either the edges or a specific portion of the trait space without affecting many other parts of the trait space.

### Find the type of metrics of interest

Once the pattern to be captured (what) has clearly been defined, we believe it becomes easier to select an appropriate metric. For example, we can follow the suggestions from Mammola et al. (2021) highlighting which kind of methodological pipeline best supports which type of metric of interest (Table 4).

### Select the metric most appropriate to your data

To describe the pattern for our specific example question, we will be interested in the size and position of the trait space. There are many different metrics that capture changes in the size and position of the trait space and their choice should depend on the following questions: do you have a lot of data or not? Do you have a lot of computational time available

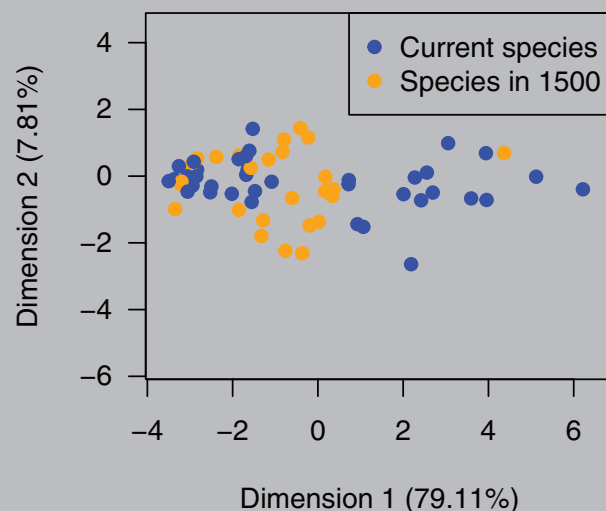


Figure 4. Example of a trait space based on the empirical data described above. 'Species in 1500' refers to now extinct species that were believed to have been extant in 1500 CE.

Table 4. Examples of questions we can ask with the different metric types. The categories are arbitrary and are taken from [Mammola et al. \(2021\)](#).

Metric type	Question	Are we interested?
Size	Is one group bigger than another one?	Yes
Divergence	Is one group more spread than another one?	Not really
Regularity	Is one group more evenly dispersed in space?	Not really
Something else	Does the position of a group change in space?	Yes

Table 5. Trait space size measured as the convex hull volume for both groups.

Group	n	Convex hull volume
Current	37	25.44
At 1500	63	60.84

Table 6. Trait space position measured as the displacement metric: the percentages are the confidence intervals of the distribution.

Group	n	Median displacement	2.5%	25%	75%	97.5%
Current	37	1.04	0.95	0.99	1.05	1.10
At 1500	63	1.12	0.85	0.94	1.19	1.34

or not? What type of data do you use (categorical, continuous, transformed)? Are you familiar with the metric (is there a software implementation you can use)? Of course, these are example questions and are not necessarily the only ones you should ask. [Mammola et al. \(2021\)](#) proposes a very handy flow chart in this regard.

Here, following our experience and the nature of the data, we will measure the absolute size of the trait space using the convex hull size and the potential displacement of the trait space using the displacement metric ([Guillerme et al. 2020b](#)). Results are displayed in [Table 5–6](#).

Note that it is also possible to test whether the metric is actually capturing the broad mechanism of interest using evolutionary simulations (e.g. by simulating evolutionary scenarios that would include species traits that would go through filtering mechanisms – [Guillerme 2024](#)) or, like in this manuscript, using stressors (e.g. by using the moms interactive package [Guillerme et al. 2020b](#)).

to random extinctions). In addition, it is worth stressing that these analyses included marine species (which constitute less than 4% of known extinctions in Hawaii, while making up 11% of the prehistoric assemblage) and different results may have been found if focusing exclusively on terrestrial birds.

More broadly, as can be seen from this empirical case study, patterns are likely to be much less clear than with simulated data due to multiple stressors acting simultaneously. For instance, extensive hunting is likely to present equalising and/or filtering pressure due to the selective hunting of larger-bodied species. In contrast, avian malaria, which is known to be an important extinction pressure for native Hawaiian birds ([Samuel et al. 2011](#)), could show signals of a null mechanism since susceptibility to the disease is not trait-dependent (at least not in relation to the dimensions we have considered here), but is related to genetics (it mainly affects passerines, but there is considerable variation in susceptibility within passerines) and distributional range ([Atkinson 2023](#)).

## Caveats

Our results may be influenced by our choice of space occupancy metrics – specifically, we focused on metrics available

in three statistical packages in R, the most common statistical language currently used in ecology ([Lai et al. 2019](#)), but many others could have been used ([Guillerme et al. 2020b](#)). Also, our results are likely affected by the use of simplified stressors designed to approximate very complex ecological and evolutionary mechanisms by simply removing a percentage of observations in trait spaces in a non-random manner. Note also that the simulated trait spaces here did not always share common characteristics with empirical trait spaces. The simulated trait spaces in 2, 4, or 8 dimensions had the same variance on all dimensions (i.e. we used the same trait simulation process for all dimensions) whereas in empirical trait spaces, commonly generated using ordination techniques (e.g. PCA, PCO, or PCoA), the dimensions have by definition a decreasing variance.

Furthermore, we only presented results based on a relatively small number of dimensions (up to 8). Although this number of dimensions is in the range of many studies in ecology (e.g. six dimensions in [Healy et al. 2019](#)), it is not uncommon to use a much greater number of dimensions, especially in palaeontology (e.g. more than 200 dimensions in [Van Den Ende et al. 2023](#)). In a higher number of dimensions (usually > 10 but this is highly variable depending on

the system under study), metric results are more difficult to scale linearly. This sometimes makes metric interpretation more chaotic (Bellman et al. 1957) in the sense that there is no direct relationship between the number of observations and dimensions that can be used to extrapolate results in higher dimensional datasets. In fact, the relationship between observations has an important effect on how our results scale with higher dimensionality (e.g. the volume of an evenly occupied space scales differently in high dimensions compared to the same space but with normally distributed observations).

Another cautionary note could be made about our decision to scale the metrics relative to a random trait space reduction (with the same percentage of reduction). This choice allowed us to simulate a null model designed to test whether the mechanisms of interest were causing the observed change in pattern (Bausman 2018). For example, if removing 80% of the edges of a distribution was distinguishable or not from removing 80% of a distribution randomly (e.g. facilitation mechanism measured as divergence using the probability density method – not distinguishable from the null when removing 80 or 60% of the data; Fig. 2). Some metrics (e.g. functional evenness using the dendrogram method) were not able to distinguish between the null mechanism and the mechanism of interest (here the mechanism simulating competition). In other words, our statistical question was ‘does metric X distinguish between removing N% of data in a biased way (the mechanism) and removing N% of data randomly (the null)?’. This definition of the null hypothesis is also appropriate for the empirical data if the question is the same. However, it is very likely that researchers will ask a more exciting or intriguing question based on these data. For example, an interesting one could be ‘are extinct species spread uniformly in the trait space of all birds in Hawaii?’. This legitimate question would thus require first a trait space (the pattern – what) and maybe some contrasting groups of interest like extinctions through time (the process – how) to answer the question (the mechanism – why). Furthermore, the scaling of the metrics paired with the levels of removal (80, 60, 40, and 20%) creates an expected artefactual similarity when comparing the random removals against the removals due to stressors: when less data are available overall, the metrics are more likely to be similar (e.g. if removing 99% of the data, we expect the random removal metric score to be nearly identical to the non-random ones).

Importantly, it should always be borne in mind that mechanisms ought of course to be more complex in real world scenarios. For example, evolutionary mechanisms can vary through time or across clades (and so patterns are often the result of multiple processes); and ecological mechanisms are often intertwined and work together to generate a pattern (e.g. facilitation + competition; Danet et al. 2024), or counteract each other by operating on the trait space in opposite directions (e.g. competition + filtering; Mammola et al. 2024).

## Conclusion

Different ecological and evolutionary processes or mechanisms do not always result in different patterns. Our results based on simulations and empirical data suggest that the same data can be interpreted differently depending on the choice of trait space occupancy metric (also known as disparity, dissimilarity, or functional diversity metrics). Different metrics are designed to capture different aspects (Guillerme et al. 2020b, Mammola et al. 2021) but also perform better or worse at their designed task depending on the data, the mechanism, or/and process of interest. Using space occupancy metrics for describing functional diversity is tricky. This is because the metrics have properties that can be counterintuitive based on the data at hand; and also because functions of ecosystems, or the organisms and niches within them, are hard to capture/understand/define. This means that measuring functional diversity is based on both the definition of the metric and the function of interest, and thus cannot be served by a one-fits-all metric. In other words, it is important to choose a disparity/dissimilarity/functional diversity metric based on the question and the data at hand rather than as a default option or solely based on a previous inspiring publication. We suggest caution when summarising patterns in observed data and propose a pipeline and tools for researchers to help understand how their pattern (often stemming from multidimensional data) can vary intuitively or not depending on the process and mechanism of interest.

*Acknowledgements* – Thanks to Andrew Beckerman, Natalie Cooper, Alain Danet, Thomas Johnson, and Gavin Thomas for comments on early versions of this manuscript.

*Funding* – TG was funded by the UKRI-NERC grant no. NE/X016781/1. SM was supported by NBFC, funded by the Italian Ministry of University and Research, PNRR, Missione 4, Componente 2, ‘Dalla ricerca all’impresa’, Investimento 1.4, Project CN00000033. MWJ was supported by NERC CENTA2 grant no. NE/S007350/1 and the University of Birmingham.

## Author contributions

**Thomas Guillerme:** Conceptualization (equal); Formal analysis (lead); Investigation (equal); Methodology (equal); Software (lead); Writing – original draft (equal); Writing – review and editing (equal). **Pedro Cardoso:** Conceptualization (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Writing – original draft (equal); Writing – review and editing (equal). **Maria Wagner Jørgensen:** Conceptualization (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Writing – original draft (equal); Writing – review and editing (equal). **Stefano Mammola:** Conceptualization (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Writing – original draft (equal); Writing – review and editing (equal). **Thomas J. Matthews:** Conceptualization (equal);

Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Writing – original draft (equal); Writing – review and editing (equal).

### Transparent peer review

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/ecog.07580>.

### Data availability statement

All code and data required to reproduce this analysis is available from the Zenodo Digital Repository: <https://doi.org/10.5281/zenodo.14852184> (Guillerme et al. 2025).

### Supporting information

The Supporting information associated with this article is available with the online version.

### References

- Anand, M. 1994. Pattern, process and mechanism – the fundamentals of scientific inquiry applied to vegetation science. – *Coenoses* 9: 81–92.
- Atkinson, C. T. 2023. Wildlife disease and health in conservation. – JHU Press.
- Barot, S. 2004. Mechanisms promoting plant coexistence: can all the proposed processes be reconciled? – *Oikos* 106: 185–192.
- Bausman, W. C. 2018. Modeling: neutral, null, and baseline. – *Philos. Sci.* 85: 594–616.
- Bellman, R., Corporation, R. and Collection, K. M. R. 1957. Dynamic programming. – Rand Corporation Research Study, Princeton Univ. Press.
- Belmaker, J. and Jetz, W. 2013. Spatial scaling of functional structure in bird and mammal assemblages. – *Am. Nat.* 181: 464–478.
- Bjorkman, A. D. et al. 2018. Plant functional trait change across a warming tundra biome. – *Nature* 562: 57–62.
- Boonman, C. C. F., Huijbregts, M. A. J., Benítez-López, A., Schipper, A. M., Thuiller, W. and Santini, L. 2022. Trait-based projections of climate change effects on global biome distributions. – *Divers. Distrib.* 28: 25–37.
- Botta-Dukát, Z. 2005. Rao's quadratic entropy as a measure of functional diversity based on multiple traits. – *J. Veg. Sci.* 16: 533–540.
- Bruno, J. F., Stachowicz, J. J. and Bertness, M. D. 2003. Inclusion of facilitation into ecological theory. – *Trends Ecol. Evol.* 18: 119–125.
- Cardoso, P., Rigal, F. and Carvalho, J. C. 2015. Bat–biodiversity assessment tools, an R package for the measurement and estimation of alpha and beta taxon, phylogenetic and functional diversity. – *Methods Ecol. Evol.* 6: 232–236.
- Cardoso, P., Guillerme, T., Mammola, S., Matthews, T. J., Rigal, F., Graco-Roza, C., Stahls, G. and Carvalho, J. C. 2024. Calculating functional diversity metrics using neighbor-joining trees. – *Ecography* 2024: e07156.
- Carmona, C. P., de Bello, F., Mason, N. W. H. and Lepš, J. 2019. Trait probability density (TPD): measuring functional diversity across scales based on TPD with R. – *Ecology* 100: e02876.
- Carvalho, J. C. and Cardoso, P. 2020. Decomposing the causes for niche differentiation between species using hypervolumes. – *Front. Ecol. Evol.* 8: 243.
- Chapin, F. S. III., Autumn, K. and Pugnaire, F. 1993. Evolution of suites of traits in response to environmental stress. – *Am. Nat.* 142: S78–S92.
- Chesson, P. 2000. Mechanisms of maintenance of species diversity. – *Annu. Rev. Ecol. Syst.* 31: 343–366.
- Chichorro, F., Urbano, F., Teixeira, D., Väre, H., Pinto, T., Brummitt, N., He, X., Hochkirch, A., Hyvönen, J., Kaila, L., Juslén, A. and Cardoso, P. 2022. Trait-based prediction of extinction risk across terrestrial taxa. – *Biol. Conserv.* 274: 109738.
- Cornwell, W. K., Schwilk, D. W. and Ackerly, D. D. 2006. A trait-based test for habitat filtering: convex hull volume. – *Ecology* 87: 1465–1471.
- Danet, A., Bautista, S., Génin, A., Beckerman, A. P., Anthelme, F. and Kéfi, S. 2024. Species diversity promotes facilitation under stressful conditions. – *Oikos* 2024: e10303.
- Dawson, S. K., Carmona, C. P., González-Suárez, M., Jönsson, M., Chichorro, F., Mallen-Cooper, M., Melero, Y., Moor, H., Simaika, J. P. and Duthie, A. B. 2021. The traits of 'trait ecologists': an analysis of the use of trait and functional trait terminology. – *Ecol. Evol.* 11: 16434–16445.
- Dushoff, J., Kain, M. P. and Bolker, B. M. 2019. I can see clearly now: reinterpreting statistical significance. – *Methods Ecol. Evol.* 10: 756–759.
- Foote, M. 1995. Morphological diversification of Paleozoic crinoids. – *Paleobiology* 21: 273–299.
- Gould, S. J. 1991. The disparity of the Burgess shale arthropod fauna and the limits of cladistic analysis: why we must strive to quantify morphospace. – *Paleobiology* 17: 411–423.
- Guillerme, T. 2018. disparity: a modular R package for measuring disparity. – *Methods Ecol. Evol.* 9: 1755–1763.
- Guillerme, T. 2024. treats: a modular r package for simulating trees and traits. – *Methods Ecol. Evol.* 15: 647–656.
- Guillerme, T. et al. 2020a. Disparities in the analysis of morphological disparity. – *Biol. Lett.* 16: 20200199.
- Guillerme, T., Puttick, M. N., Marcy, A. E. and Weisbecker, V. 2020b. Shifting spaces: which disparity or dissimilarity measurement best summarize occupancy in multidimensional spaces? – *Ecol. Evol.* 10: 7261–7275.
- Guillerme, T., Bright, J. A., Cooney, C. R., Hughes, E. C., Varley, Z. K., Cooper, N., Beckerman, A. P. and Thomas, G. H. 2023. Innovation and elaboration on the avian tree of life. – *Sci. Adv.* 9: eadg1641.
- Guillerme, T., Cardoso, P., Jørgensen, M. W., Mammola, S. and Matthews, T. J. 2025. Data from: The what, how, and why of trait-based analyses in ecology. – Zenodo Digital Repository, <https://doi.org/10.5281/zenodo.14852184>.
- Hardin, G. 1960. The competitive exclusion principle: an idea that took a century to be born has implications in ecology, economics and genetics. – *Science* 131: 1292–1297.
- Healy, K., Ezard, T. H., Jones, O. R., Salguero-Gómez, R. and Buckley, Y. M. 2019. Animal life history is shaped by the pace of life and the distribution of age-specific mortality and reproduction. – *Nat. Ecol. Evol.* 3: 1217–1224.
- Jarzyna, M. A. and Jetz, W. 2018. Taxonomic and functional diversity change is scale dependent. – *Nat. Commun.* 9: 2565.
- Lai, J., Lortie, C. J., Muenchen, R. A., Yang, J. and Ma, K. 2019. Evaluating the popularity of R in ecology. – *Ecosphere* 10: e02567.

- Laliberté, E. and Legendre, P. 2010. A distance-based framework for measuring functional diversity from multiple traits. – *Ecol. Evol.* 1: 299–305.
- Lamanna, C. et al. 2014. Functional trait space and the latitudinal diversity gradient. – *Proc. Natl Acad. Sci. USA* 111: 13745–13750.
- Luza, A. L., Barneche, D. R., Cordeiro, C. A., Dambros, C. S., Ferreira, C. E. L., Floeter, S. R., Giglio, V. J., Luiz, O. J., Mendes, T. C., Picoletto, V. A. P., Quimbayo, J. P., Silva, F. C., Waechter, L., Longo, G. O. and Bender, M. G. 2023. Going across taxa in functional ecology: review and perspectives of an emerging field. – *Funct. Ecol.* 12: 3091–3110.
- MacArthur, R. and Levins, R. 1967. The limiting similarity, convergence, and divergence of coexisting species. – *Am. Nat.* 101: 377–385.
- Mammola, S. and Cardoso, P. 2020. Functional diversity metrics using kernel density n-dimensional hypervolumes. – *Methods Ecol. Evol.* 11: 986–995.
- Mammola, S., Carmona, C. P., Guillaume, T. and Cardoso, P. 2021. Concepts and applications in functional diversity. – *Funct. Ecol.* 35: 1869–1885.
- Mammola, S., Graco-Roza, C., Ballarin, F., Hesselberg, T., Isaia, M., Lunghi, E., Mouron, S., Pavlek, M., Tolve, M. and Cardoso, P. 2024. Functional convergence underground? The scale-dependency of community assembly processes in European cave spiders. – *Global Ecol. Biogeogr.* 2024: e13840.
- Martin, A. R. and Isaac, M. E. 2015. Plant functional traits in agroecosystems: a blueprint for research. – *J. Appl. Ecol.* 52: 1425–1435.
- Martínez, A., García-Gómez, G., García-Herrero, Á., Sánchez, N., Pardos, F., Izquierdo-Muñoz, A., Fontaneto, D. and Mammola, S. 2021. Habitat differences filter functional diversity of low dispersive microscopic animals (Acari, Halacaridae). – *Hydrobiologia* 848: 2681–2698.
- Matthews, T. J., Wayman, J. P., Cardoso, P., Sayol, F., Hume, J. P., Ulrich, W., Tobias, J. A., Soares, F. C., Thébaud, C., Martin, T. E. and Triantis, K. A. 2022. Threatened and extinct island endemic birds of the world: distribution, threats and functional diversity. – *J. Biogeogr.* 49: 1920–1940.
- Matthews, T. J. et al. 2023. A global analysis of avian island diversity–area relationships in the Anthropocene. – *Ecol. Lett.* 26: 965–982.
- McGill, B. J., Enquist, B. J., Weiher, E. and Westoby, M. 2006. Rebuilding community ecology from functional traits. – *Trends Ecol. Evol.* 21: 178–185.
- McLean, M., Stuart-Smith, R. D., Villéger, S., Auber, A., Edgar, G. J., MacNeil, M. A., Loiseau, N., Leprieur, F. and Mouillot, D. 2021. Trait similarity in reef fish faunas across the world's oceans. – *Proc. Natl Acad. Sci. USA* 118: e2012318118.
- Novack-Gottshall, P. M. 2016a. General models of ecological diversification. I. Conceptual synthesis. – *Paleobiology* 42: 185–208.
- Novack-Gottshall, P. M. 2016b. General models of ecological diversification. ii. Simulations and empirical applications. – *Paleobiology* 42: 209–239.
- Palacio, F. X., Callaghan, C. T., Cardoso, P., Hudgins, E. J., Jarzyna, M. A., Ottaviani, G., Riva, F., Graco-Roza, C., Shirey, V. and Mammola, S. 2022. A protocol for reproducible functional diversity analyses. – *Ecography* 2022: e06287.
- Pavoine, S. and Bonsall, M. B. 2011. Measuring biodiversity to explain community assembly: a unified approach. – *Biol. Rev.* 86: 792–812.
- Petchey, O. L. and Gaston, K. J. 2002. Functional diversity (FD), species richness and community composition. – *Ecol. Lett.* 5: 402–411.
- Pigot, A. L., Sheard, C., Miller, E. T., Bregman, T. P., Freeman, B. G., Roll, U., Seddon, N., Trisos, C. H., Weeks, B. C. and Tobias, J. A. 2020. Macroevolutionary convergence connects morphological form to ecological function in birds. – *Nat. Ecol. Evol.* 4: 230–239.
- R Core Team. 2024. R: a language and environment for statistical computing. – R Foundation for Statistical Computing.
- Raup, D. M. 1961. The geometry of coiling in gastropods. – *Proc. Natl Acad. Sci. USA* 47: 602–609.
- Samuel, M. D., Hobbelen, P. H. F., DeCastro, F., Ahumada, J. A., LaPointe, D. A., Atkinson, C. T., Woodworth, B. L., Hart, P. J. and Duffy, D. C. 2011. The dynamics, transmission, and population impacts of avian malaria in native Hawaiian birds: a modeling approach. – *Ecol. Appl.* 21: 2960–2973.
- Sayol, F., Cooke, R. S. C., Pigot, A. L., Blackburn, T. M., Tobias, J. A., Steinbauer, M. J., Antonelli, A. and Faurby, S. 2021. Loss of functional diversity through anthropogenic extinctions of island birds is not offset by biotic invasions. – *Sci. Adv.* 7: eabj5790.
- Sheard, C., Neate-Clegg, M. H. C., Alioravainen, N., Jones, S. E. I., Vincent, C., MacGregor, H. E. A., Bregman, T. P., Claramunt, S. and Tobias, J. A. 2020. Ecological drivers of global gradients in avian dispersal inferred from wing morphology. – *Nat. Commun.* 11: 2463.
- Silverman, B. 1986. Density estimation for statistics and data analysis. – Chapman & Hall/CRC Monographs on Statistics & Applied Probability.
- Tobias, J. A. et al. 2022. Avonet: morphological, ecological and geographical data for all birds. – *Ecol. Lett.* 25: 581–597.
- Tucker, C. M., Cadotte, M. W., Carvalho, S. B., Davies, T. J., Ferrier, S., Fritz, S. A., Grenyer, R., Helmus, M. R., Jin, L. S., Mooers, A. O., Pavoine, S., Purschke, O., Redding, D. W., Rosauer, D. F., Winter, M. and Mazel, F. 2017. A guide to phylogenetic metrics for conservation, community ecology and macroecology. – *Biol. Rev.* 92: 698–715.
- Van Den Ende, C., Puttick, M. N., Urrutia, A. O. and Wills, M. A. 2023. Why should we compare morphological and molecular disparity? – *Methods Ecol. Evol.* 14: 2390–2410.
- Villéger, S., Mason, N. W. H. and Mouillot, D. 2008. New multidimensional functional diversity indices for a multifaceted framework in functional ecology. – *Ecology* 89: 2290–2301.
- Violle, C., Navas, M.-L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. and Garnier, E. 2007. Let the concept of trait be functional! – *Oikos* 116: 882–892.
- Violle, C., Reich, P. B., Pacala, S. W., Enquist, B. J. and Kattge, J. 2014. The emergence and promise of functional biogeography. – *Proc. Natl Acad. Sci. USA* 111: 13690–13696.
- Walther, M. and Hume, J. 2022. Extinct birds of Hawai'i. – Mutual Publishing.