ECOGRAPHY

Research article

A protocol for reproducible functional diversity analyses

Facundo X. Palacio, Corey T. Callaghan, Pedro Cardoso, Emma J. Hudgins, Marta A. Jarzyna, Gianluigi Ottaviani, Federico Riva, Caio Graco-Roza, Vaughn Shirey and Stefano Mammola

F. X. Palacio, División Zoología Vertebrados, Museo de La Plata, Univ. Nacional de La Plata, La Plata, Argentina and Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), La Plata, Argentina. – C. T. Callaghan (https://orcid.org/0000-0003-0415-2709), German Centre for Integrative Biodiversity Research (iDiv) – Leipzig, Halle, Jena, Leipzig, Germany. – P. Cardoso (https://orcid.org/0000-0001-8119-9960) and S. Mammola (https://orcid.org/0000-0002-4471-9055) ✉ (stefano.mammola@cnr.it), LIBRe–Laboratory for Integrative Biodiversity Research, Finnish Museum of Natural History (LUOMUS), Univ. of Helsinki, Helsinki, Finland. – E. J. Hudgins (https://orcid.org/0000-0002-8402-5111) and F. Riva (https://orcid.org/0000-0002-1724-4293), Dept of Biology, Carleton Univ., Ottawa, ON, Canada. – M. A. Jarzyna (https://orcid.org/0000-0002-6734-0566), Dept of Evolution, Ecology and Organismal Biology, The Ohio State Univ., Columbus, OH, USA and Translational Data Analytics Inst., The Ohio State Univ., Columbus, OH, USA. – G. Ottaviani (https://orcid.org/0000-0003-3027-4638), Inst. of Botany, The Czech Academy of Sciences, Třeboň, Czech Republic. – C. Graco-Roza (https://orcid.org/0000-0002-0353-9154), Dept of Geosciences and Geography, Univ. of Helsinki, Helsinki, Finland. – V. Shirey (https://orcid.org/0000-0002-3589-9699), Dept of Biology, Georgetown Univ., Washington, DC, USA. SM also at: Molecular Ecology Group (MEG), Water Research Inst. (IRSA), National Research Council of Italy (CNR), Verbania Pallanza, Italy.

Ecography
2022: e06287
doi: 10.1111/ecog.06287
Subject Editor: Dominique Gravel
Editor-in-Chief: Miguel Araújo
Accepted 1 August 2022

The widespread use of species traits in basic and applied ecology, conservation and biogeography has led to an exponential increase in functional diversity analyses, with > 10 000 papers published in 2010–2020, and > 1800 papers only in 2021. This interest is reflected in the development of a multitude of theoretical and methodological frameworks for calculating functional diversity, making it challenging to navigate the myriads of options and to report detailed accounts of trait-based analyses. Therefore, the discipline of trait-based ecology would benefit from the existence of a general guideline for standard reporting and good practices for analyses. We devise an eight-step protocol to guide researchers in conducting and reporting functional diversity analyses, with the overarching goal of increasing reproducibility, transparency and comparability across studies. The protocol is based on: 1) identification of a research question; 2) a sampling scheme and a study design; 3–4) assemblage of data matrices; 5) data exploration and preprocessing; 6) functional diversity computation; 7) model fitting, evaluation and interpretation; and 8) data, metadata and code provision. Throughout the protocol, we provide information on how to best select research questions, study designs, trait data, compute functional diversity, interpret results and discuss ways to ensure reproducibility in reporting results. To facilitate the implementation of this template, we further develop an interactive web-based application (*stepFD*) in the form of a checklist workflow, detailing all the steps of the protocol and allowing the user to produce a final 'reproducibility report' to upload alongside the published paper. A thorough and transparent reporting of functional diversity analyses ensures that ecologists can incorporate others' findings into meta-analyses, the shared data can be

NORDIC SOCIETY OIKOS

www.ecography.org

integrated into larger databases for consensus analyses, and available code can be reused by other researchers. All these elements are key to pushing forward this vibrant and fast-growing field of research.

Keywords: biological diversity, ecosystem functioning, open science, replicability, reproducibility, standardised protocols, trait-based ecology

Introduction

Failure to reproduce many results in the published literature is causing discussions among scientists about poor research practices (Baker 2016, Fanelli 2018). A lack of reproducibility (Glossary) hinders our ability to corroborate or falsify results, and is often associated with incomplete reporting of experimental protocols and pipelines (Munafò et al. 2017), limited data and code sharing (Tenopir et al. 2011, Culina et al. 2020) and misuse of statistics and analyses (e.g. cherry picking statistically significant results, p-hacking, hypothesising after the results are known; Fraser et al. 2018). Like in other scientific disciplines, ecologists do not evade such practices (Borregaard and Hart 2016). For example, reproducibility has become highly relevant in field-based ecological studies, due to the impossibility of exactly repeating the results obtained via observational data (e.g. along environmental gradients), where accounting for abiotic and biotic factors is more challenging than in controlled laboratory conditions (Ellison 2010, Powers and Hampton 2019). Concerns over transparent practices in ecology (Fidler et al. 2017, Fraser et al. 2018, Culina et al. 2020, Eckert et al. 2020) have prompted the development of protocols to enhance and achieve best standards in data acquisition, including pipelines and protocols for conducting regression-type analyses (Zuur and Ieno 2016), modelling species distributions (Araújo et al. 2019, Feng et al. 2019, Zurell et al. 2020), performing phenotypic selection analyses in evolutionary ecology (Palacio et al. 2019b) and collecting trait data (Cornelissen et al. 2003, Moretti et al. 2017, Klimešová et al. 2019).

Despite this progress, discussions about reproducibility are still incipient in trait-based ecology. Trait-based studies have increased exponentially in the last 20 years (Fig. 1), advancing our understanding of the impact of global change on biodiversity (Newbold et al. 2020), ecological resilience (Pausas et al. 2016) and determinants of assembly rules at different spatial and temporal scales (Mouillot et al. 2021). As a result, estimating functional (or trait) diversity (Glossary) has emerged as one of the core constructs in modern ecology (McGill et al. 2006). This broad interest has prompted the development of a myriad of methods and metrics (Schleuter et al. 2010, Pavoine and Bonsall 2011, Mammola et al. 2021), making it difficult to select appropriate methods for answering specific ecological questions and to keep track of new concepts and approaches. Given that conclusions drawn from any given study are sensitive to how data are collected, handled (e.g. methodological choices preceding functional diversity computation) and analysed (Lavorel et al. 2007, Maire et al. 2015, Perronne et al. 2017), there is urgent need for addressing reproducibility practices in trait-based ecology and developing a standard for reporting study design and analyses.

To this end, we have devised a general roadmap for transparent reporting of all the steps typical of any trait-based study. We developed an eight-step protocol from study inception and design to data, metadata and code reporting (Fig. 2). We suggest that trait-based studies start with the conceptualisation of an ecological question, generally ingrained in a theoretical hypothesis-driven framework (Step 1). A clear ecological rationale then informs an appropriate experimental design (Step 2). Next, occurrence (Step 3) and trait (Step 4) data – the raw material of any trait analysis – are collected. Data exploration (Step 5) precedes the core of the analysis with functional traits (Step 6) and the validation, interpretation and reporting of results (Step 7). The last step (Step 8) encompasses data and code sharing to ensure the reproducibility of the entire pipeline. The protocol provided here is geared primarily toward scientists who are entering the discipline of trait-based ecology and have little to moderate experience with functional diversity analysis. More experienced users, however, might also find the protocol useful, particularly the elements pertaining to data analysis (Step 6), validation and interpretation (Step 7) and transparent reporting of the study (Step 8 and the 'reproducibility report' developed via a Shiny web application).

Step 1. Identify an appropriate research question

Most scientific study begins with a question or hypothesis, and it is therefore critical to establish a salient and feasible one prior to collecting data. Because resources are often limited, one should ensure that the question addressed has theoretical and/or applied relevance, while being methodologically and logistically feasible. Researchers must therefore first evaluate whether examining functional diversity might provide more in-depth (or complementary) insights into the question of interest than other approaches (e.g. taxonomic or phylogenetic).

Then, once it has been established that functional diversity is relevant, one can follow two main scientific approaches: the hypothetico-deductive (formulating hypotheses first, and then testing these hypotheses by collecting data) or inductive (collecting empirical observations first, and then generating potential explanations of the patterns observed) paradigms (Mentis 1988). Hypothetico-deductive approaches are based

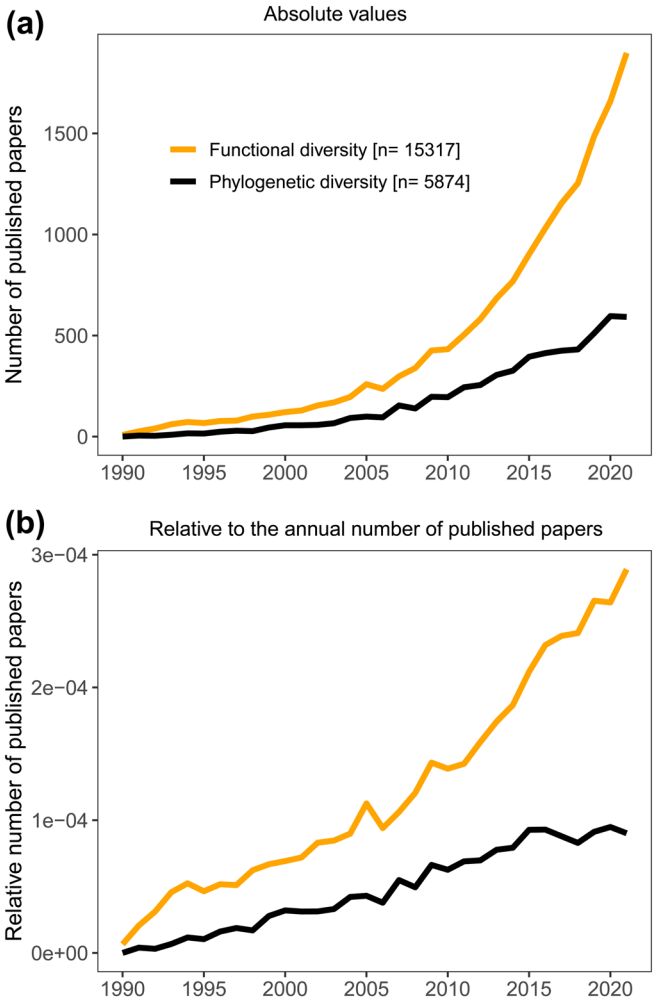

Figure 1. (A) Annual (1990–2021) number of published papers using the term 'functional diversity' compared to 'phylogenetic diversity'. (B) Number of papers using the two terms relativized to the total annual number of published papers, to account for the general growth in scientific literature volume in recent years (Landhuis 2016). The number of papers was sourced from the Web of Science (Clarivate Analytics) on 28 December 2021, using the queries: TS = 'functional diversity' and TS = 'phylogenetic diversity'. The total number of papers published each year is based on the Dimensions database, accessed on 12 January 2021.

on the idea that one can untangle the complexity of natural systems by testing alternative hypotheses (see e.g. strong inference; Platt 1964). Many have argued that a hypothetico-deductive scheme has led to more advancements in scientific understanding (Platt 1964, Betts et al. 2021), but the inductive scheme also plays important roles, especially in creating foundational knowledge (Mentis 1988). The choice between hypothetico-deductive and inductive frameworks is a matter of philosophical preferences that is generally also constrained by the question and the scale of the study. For instance, plants and microorganisms are relatively easy to experimentally manipulate in terms of their abundance, trait values or simplified abiotic and biotic conditions (typically at small spatio-temporal scales), and thus may allow easier implementation of the hypothetico-deductive scheme. Conversely, inductive approaches are especially relevant in studies involving monitoring through space and time, for example to assess how disturbance or protection affect functional diversity from past to current situations. Under these circumstances, predictive power can be more important than ecological interpretation of a model, or than understanding the mechanisms underlying a pattern of interest (Currie 2019, Betts et al. 2021). Philosophical differences between the two approaches are discussed extensively elsewhere (Mentis 1988, Betts et al. 2021). Here, we simply suggest that studies of functional diversity can be useful in both views, and call for consideration of these general conceptual aspects while conceiving study designs.

Step 2. Identify an appropriate study design

The choice of the study design – experimental, observational, simulation or meta-analytical – should be dictated by the research question(s) (Step 1). Experimental studies allow controlling for major confounding factors inherent to natural settings, but often represent simplified systems. For instance, an experiment can isolate the role of biotic interactions (e.g. competition, facilitation) in community assembly by manipulating community trait composition or environmental conditions, typically at very fine scales. Observational studies facilitate insights into ecological patterns, but their ability to disentangle the mechanisms underlying them is often limited (de Bello et al. 2012, Spasojevic and Suding 2012, Cadotte and Tucker 2017). In parallel, simulations can be used to link patterns revealed from observational studies with putative processes to evaluate conditions in which a given process might result in an observed pattern. Simulations can also pinpoint numerical properties and statistical artefacts, which are especially important in functional diversity studies where subjective choices, e.g. on the number, type and measure of traits, are routinely made (McPherson et al. 2018; see Step 4). Finally, meta-analyses can increase the generalisability of individual studies, and may facilitate the resolution of inconclusive or conflictive results (Greenop et al. 2018, Woodcock et al. 2019, Matuoka et al. 2020).

The study design should be considered in the context of data availability and limitations (Steps 3 and 4). Available databases vary in relation to their spatial coverage and extent, with spatio-temporal resolution typically decreasing with spatial extent (Hulbert and Jetz 2007). Occurrence and trait data sources (opportunistic, historical or collected/experiment) are a primary consideration when designing a study. At the same time, open source datasets (Supporting information), community science datasets (Callaghan et al. 2021) and museum/herbarium collections are becoming increasingly important in trait-based ecology (Perez et al. 2020). The scale of analysis will also determine whether a trait data source is appropriate for use. At small scales (e.g. when setting up an experiment), global databases may not

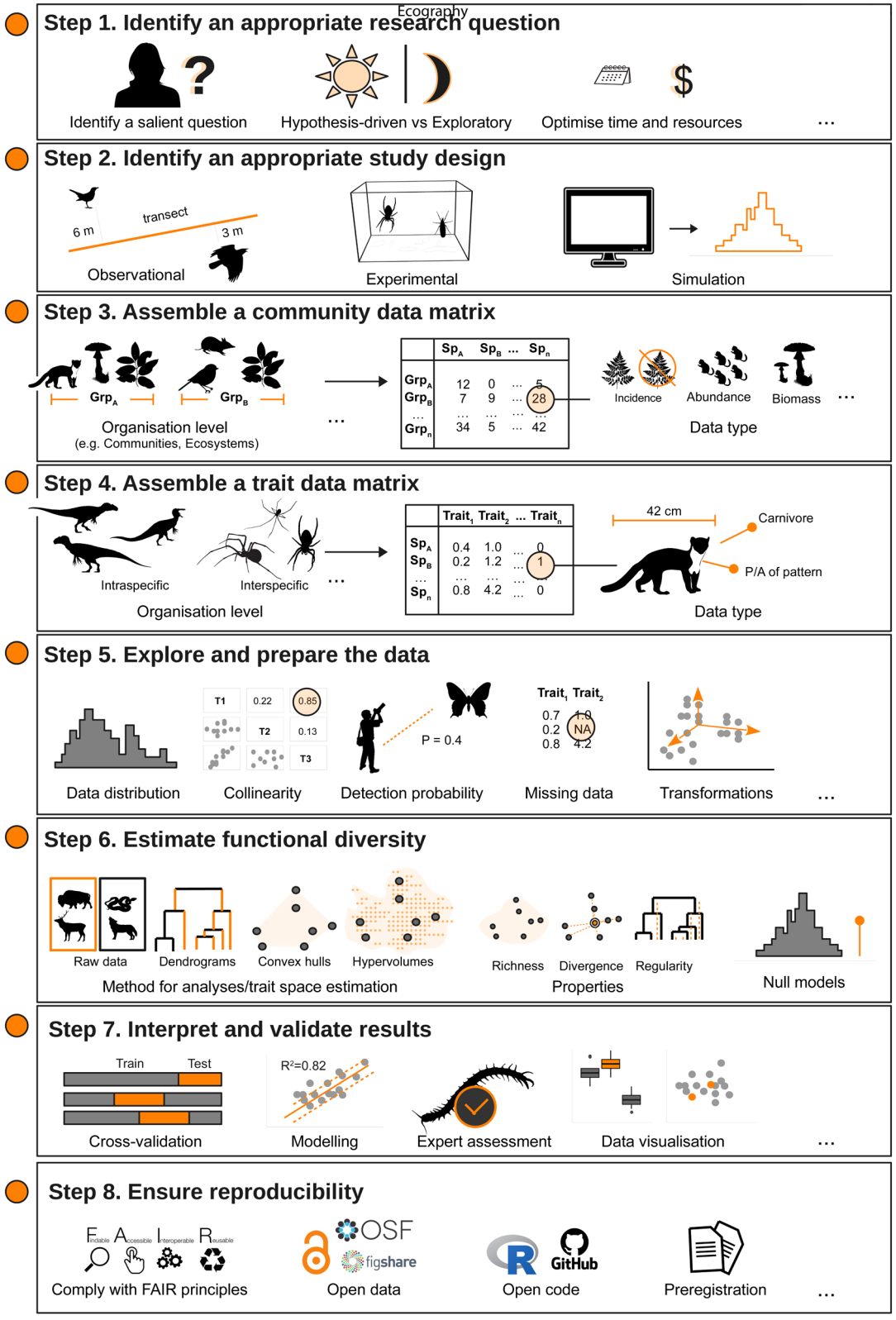

Figure 2. Workflow of the eight-step protocol proposed in this study. Animal silhouettes retrieved from Phylopics – with open licence.

be appropriate to capture functional diversity and one's own measurements are preferable. Conversely, global datasets with coarser trait resolution are most useful when assessing macroecological patterns in functional diversity across large geographic extents (Violle et al. 2014), especially if focusing on interspecific differences (intraspecific variability is generally not included in global datasets).

If the functional diversity study is based on the collection of one's own occurrence and trait data, then the identification of an appropriate sampling protocol is the next crucial step. This should be primarily driven by the research question (Step 1), the scale of the focal ecological phenomenon (McGill 2010), and the level of organisation at which functional diversity will be assessed (e.g. individuals within a population, species forming a community, communities in a regional pool; Violle et al. 2014). A more comprehensive overview of different approaches in experimental design (Scheiner and Gurevitch 2001), and of their importance (Christie et al. 2019) is outside the focus of this manuscript.

Step 3. Assemble a community data matrix

Once the data has been collected (Step 2), these need to be formatted in a meaningful way to explore functional diversity. Observations are organised in a community data matrix **C** holding data on the occurrence of species (in a broad sense, here referred to any type of data describing incidence, abundance, biomass or coverage of individuals; see below). Note that we used here the term 'community matrix' being the most typical level of organisation at which functional diversity is measured; but this matrix could hold data on individuals within species, communities within ecosystems or any other grouping level, and be analysed using similar principles. In the most common case, this is a matrix of S rows × n columns, where rows ($i = 1, 2, \ldots, S$) represent sampling units (e.g. sites, plots, transects) and columns ($j = 1, 2, \ldots, n$) represent taxonomic entities of interest (typically species, but also individuals or higher taxonomic ranks) found within each sampling unit. This basic matrix can be expanded to a set of temporal replicates or a set of individuals when accounting for intraspecific variation. This can be achieved by including additional rows accounting for different times at the same site or different individuals of the same species. According to the method used, one may need to include additional columns describing the grouping level (e.g. site, species), as well as columns in the trait data matrix matching rows in the community matrix (see more details in Step 4). In describing the matrix **C**, one should specify taxonomic resolution, sample size (i.e. number of sampling units, temporal replicates), number of recorded taxa and sampling effort.

Data on the occurrence of species may take multiple forms with different ecological interpretations, which should be clarified. Incidence (presence/absence) and abundance (number of individuals) data have historically been most commonly used in community ecology. Nevertheless, presence-only data or model-based estimates of species incidence/abundance have also been used. Presence-only data (usually derived from online databases or historical museum records) are probably not suited for estimating functional diversity, because they introduce multiple sources of biases (e.g. different collection methods and sampling effort). Moreover, the community matrix derived from presence-only data assumes zeroes as true absences. By contrast, occupancy modelling (Box 1) accounts for differences in sampling effort, overcoming many of the issues of presence-only data (Shirey et al. 2021). The community matrix in this case is composed of occupancy probabilities (Box 1).

Box 1. Species detectability and functional diversity estimation

Perfect detection of organisms is rare, often resulting in false species absences or the underestimation of population sizes and biodiversity. Estimates of functional diversity can be disproportionately affected by such 'missed detections', because species detectability is often linked to their functional distinctiveness or certain trait characteristics (including trait resolution; Jarzyna and Jetz 2016), suggesting that specific traits might be underestimated or missed altogether during data collection (Roth et al. 2018, Palacio et al. 2020). The magnitude and the direction of the impact of species detectability on FD estimates will depend on several factors (Jarzyna and Jetz 2016), including 1) the proportion of undetected species at a site and their traits, 2) the size of the regional species pool, 3) the spatial scale at which data are collected and FD is estimated and 4) how species detectability varies along spatial and environmental gradients (Jarzyna and Jetz 2016, Palacio et al. 2020).

Recent advances in statistical modelling allow accounting for species' detectability. Specifically, multispecies occupancy (Iknayan et al. 2014, Denes et al. 2015) and N-mixture (Gomez et al. 2018) models allow for estimation of the 'true' probability of each species occurrence or their detection-corrected abundance, which can then be incorporated into functional diversity estimates (Jarzyna and Jetz 2016, Palacio et al. 2020). Multispecies occupancy and N-mixture models can be fitted in either a frequentist or a Bayesian framework (Devarajan et al. 2020). If models are fitted in a Bayesian framework, it is advised to report algorithmic details for initial values for parameter estimation, prior distributions and a summary of posterior estimates (e.g. occurrence and detection probabilities). Depending on the methodology used, details such as the number of Markov chains and iterations per chain, burn-in, the thinning parameter convergence evaluation may need to be reported as well.

As another alternative to deal with this type of data, species distribution models can be fitted to obtain suitability values and used as inputs of the community matrix (Andrew et al. 2021). It should be noted that both models are of recent application, so there are no general guidelines on which approach is better to handle presence-only data. Other types of data, such as biomass and percent cover in sessile organisms, are often treated as abundance proxies or transformed into incidence data (Riva et al. 2020).

All these types of data can come from different sources. Besides laboratory/field experiments and traditional observations, rapid progression in monitoring technologies (e.g. remote sensing, acoustic sensors, camera traps, environmental DNA, metabarcoding) has enabled ecologists to automate extraction of massive amounts of biodiversity data from different environmental media (e.g. water, soil, air), and identify taxa associated with the environment with high accuracy (Tosa et al. 2021). Whilst promising, the use of these data sources is still at an incipient state in trait-based ecology (Gasc et al. 2013, Schneider et al. 2017, Aglieri et al. 2020, Sigsgaard et al. 2020). Given method-specific technical limitations (e.g. amplification of a large proportion of nontarget sequences and degradation time of DNA), we suggest always reporting whether sampling effort has been adequate to capture taxonomic diversity – e.g. through rarefaction techniques (Roswell et al. 2021).

Step 4. Assemble a trait data matrix

The second key element of any functional diversity analysis is the use of species traits. These include a variety of morphological, behavioural, physiological, anatomical, biochemical or phenological attributes that have the potential to impact the individual's fitness (Violle et al. 2007). We note that there is an ongoing debate on terminology in trait-based ecology (Volaire et al. 2020, Dawson et al. 2021, Sobral 2021), which is beyond the scope of our paper. Regardless of the definition used, traits provide the raw material to build the trait data matrix **T**. This is a matrix of R rows × P columns where rows ($i = 1, 2, \ldots, R$) represent the taxonomic entities of interest (univocally corresponding to the n columns in the **C** matrix when there is one trait value per taxonomic entity), and columns ($j = 1, 2, \ldots, P$) represent traits.

Trait data can be measured directly (e.g. in the field/laboratory or from museum specimens), extracted from different sources (e.g. peer-reviewed literature, field guides, online databases; Supporting information) or a combination of the above. Trait resolution (Glossary) should be carefully considered, particularly when different data sources are combined, as differences in resolution may confound ecological patterns and bias inference (Cordlandwehr et al. 2013, Palacio et al. 2019a, Kohli and Jarzyna 2021). For instance, using coarse-resolution categories as a substitute for continuous traits inevitably masks trait variability, inflating functional redundancy and decreasing functional distances among species, and, consequently, perceived functional diversity. Importantly, trait resolution is also expected to impact the estimation of underlying deterministic processes, which are typically inferred from patterns of functional divergence and convergence. For example, while the detection of a true underlying structure for communities that are functionally convergent will not be affected by employing coarse-resolution traits, the ability to detect true functional divergence is likely to be significantly compromised (see Fig. 1 in Kohli and Jarzyna 2021). However, coarse-resolution categorical traits are the most widely available and used in functional diversity analysis for most taxonomic groups. We thus advise that researchers report trait resolution of traits that are not intrinsically categorical by indicating how many categories were used to split the trait.

Choosing how many traits to include is also not trivial. For instance, there might be trade-offs between using a low number of traits yielding limited variability to properly estimate functional diversity, or using a high number of traits yielding too many unique combinations of trait values (in the most extreme case, functional diversity may equal species richness; Petchey and Gaston 2002). We refer to Step 5 for further insights.

Researchers should also keep in mind that the same trait might represent different processes and functions for different taxa or in different contexts. As an example, larger body size might imply a limitation of resource availability for animals, but may allow plants to outcompete others in the search for light. Similarly, the same function might be represented by different traits in different taxa. For example, dispersal ability is represented by the ratio between wing and body size and shape for many insects (Lancaster and Downes 2017), the ability and propensity to balloon for spiders (Bonte et al. 2003), the seed size and dispersal modes for aquatic plants (de Jager et al. 2019), and the tendency to be entrained in long-distance transport vectors in invasive species (Hastings et al. 2005). Therefore, if the aim of the study is to analyse how a certain factor impacts on a given ecosystem function in different taxa, then one should select traits capturing the same function.

The ecological rationale for which traits are selected in the analysis is also critical and should be carefully described, along with the specific function(s) the trait is able to represent (Weiher et al. 1999, Luck et al. 2012). Most functional diversity studies rely on species' mean trait values – i.e. averaged across trait measurements collected from multiple individuals per species ('mean field approach' sensu Violle et al. 2012). This relies on the assumption that among-species trait variation largely exceeds intraspecific trait variation. However, growing evidence shows that intraspecific variation can be substantial and affects different ecological processes (Albert et al. 2011, Palacio et al. 2019b, Gentile et al. 2021, Wong and Carmona 2021). For instance, trait values of species may vary along an environmental gradient due to phenotypic plasticity and/or local adaptation increasing intraspecific trait variation (Günter et al. 2019). As a result, two communities with the same species composition may have different trait distributions and thus different functional

diversity. Our protocol therefore calls for a clear statement whether trait data are described by measurements collected from several individuals and averaged at the species level, or if intraspecific variation has been taken into account and at which organisational level (e.g. site, species, populations, individuals, specific organs).

The matrix **T** can easily accommodate multiple measurements per trait, for instance, when intraspecific trait variation is of interest. Some methods, such as functional dendrograms, require both *S* and *R* to be equal to the total number of trait measurements such that **C** and **T** are conformable matrices for multiplication (in other words, these matrices are $\mathbf{C}_{N\times S}$ and $\mathbf{T}_{S\times P}$, termed 'site × individual' and 'individual × trait' matrices, respectively) and return an individual-based functional dendrogram as a functional space. Other methods, such as the trait probability density (TPD) framework, do not require initial matrix conformability because the **T** matrix is preprocessed before being combined with the information in the **C** matrix (see detailed explanation in Carmona et al. 2016). Likewise, in the case of hypervolumes, functional trait spaces can be obtained via the union of individual-level functional hypervolumes (see details in Mammola and Cardoso 2020, Graco-Roza et al. 2022).

We ultimately recommend detailing the traits used, their nature (e.g. indicating their possible states or range values, and the ontogenetic stages of the sampled individuals), and their hypothesised ecological function(s). The methods should also contain all relevant information on trait data sources. If trait data are retrieved from online databases, then information on version and access date should be reported.

Step 5. Explore and prepare the data

Data exploration is perhaps one of the most informative, yet often overlooked, steps of analysing an ecological dataset (Zuur et al. 2010). When inspecting the community data matrix (Step 3), one has to carefully check for the existence and potential causes of zero-inflation in occurrence data (these can be true zeros or an artefact due to, e.g. imperfect detection, species misidentification or poor sampling design; Roth et al. 2018, Blasco-Moreno et al. 2019), dependency structures (e.g. spatio-temporal autocorrelation) and potential problems due to uneven spatio-temporal sampling effort (Walker et al. 2008, Ricotta et al. 2012). Trait data (Step 4) are often a mixture of numerical, ordered, fuzzy, and/or categorical variables that should be examined for correlation. Trait data can also be characterised by unbalanced levels in categorical traits, outliers in continuous traits and missing data, all of which might condition the trait space and functional diversity estimation (Step 6).

To ensure data quality and integrity in both occurrence and trait data, as a general pipeline, we recommend to:

1. Plot the community data matrix (e.g. heatmaps) to assess the prevalence of zeroes (Box 1).
2. Check species sampling coverage (e.g. rarefaction).
3. Plot the distribution of continuous traits (e.g. using histograms, density plots, Cleveland dot plots, correlograms or boxplots) to check for outliers. Plot categorical traits (e.g. with barplots) to check the balance of levels in fuzzy and categorical variables.
4. Evaluate multicollinearity among continuous traits (e.g. with scatterplots, pairwise correlations) and associations between continuous and categorical traits (e.g. with boxplots).
5. Identify missing trait data (e.g. with barplots or heatmaps).

These steps provide a better understanding into the nature of, and the issues inherent to the data, and thus allow making informed decisions on how to best approach the analysis. Depending on the outcome of initial data exploration, researchers might need to decide: 1) whether statistical

Box 2. Missing data and data imputation

Because encountering species in the field and measuring relevant traits can be difficult, trait matrices often contain missing data, which can be randomly distributed or not (Nakagawa and Freckleton 2008). Missing data need to be dealt with in order to compute virtually any method for estimating functional diversity. Three main options are available: 1) omit the individuals/species for which trait data are missing, 2) impute the missing trait data and 3) convert the trait matrix using a distance measure that allows the presence of missing data (e.g. Gower distance; de Bello et al. 2021b). If omission is the selected strategy, the consequences of removing observations linked to missing trait data should be understood and discussed. Alternatively, one might use imputation methods (Penone et al. 2014, Taugourdeau et al. 2014, Johnson et al. 2021), which are roughly based on two strategies: 1) replacing the missing value with a systematically chosen value from the phylogenetically/functionally most similar species; or 2) predicting the missing trait value, e.g. based on linear models (potentially including a phylogenetic covariance structure; Johnson et al. 2021) or principal component analysis (Podani et al. 2021), where traits are estimated as a function of other variables. Depending on whether the missing data are random or not, different algorithms should be considered for the imputation (Wulff and Jeppesen 2017). Finally, some simply use 'average imputation', calculating the mean or median of the values for that trait based on all the non-missing observations. This has the advantage of keeping the same mean and the same sample size but many disadvantages, and thus we discourage this strategy (Taugourdeau et al. 2014; see also Denny 2017 for a theoretical discussion).

Box 3. Selecting the optimal trait space dimensionality

How many traits should be used in the analysis? Although it might sound like a trivial question, there is still no consensus. The optimal number of traits (or functional axes when using dimensionality reduction techniques) is often system-, taxon- and method-dependent, with no guidelines that will work in all cases. On top of ecological and methodological considerations, too many dimensions may introduce multi-collinearity in the analysis, degrading predictive power (Dormann et al. 2013) and increasing the likelihood of type II statistical error (Zuur et al. 2010). Therefore, we recommend researchers to evaluate dimensionality of the analysis on a case-by-case basis, eventually comparing the performance of the analysis across different numbers of dimensions. Recently, Mouillot et al. (2021) showed that, in most cases, between 3 and 6 functional axes should be enough to accurately describe the matrix **T** without significant information loss. Yet, there is considerable variation among taxonomic groups (Díaz et al. 2016, Pigot et al. 2020) and this inference was based on a single method for estimating functional diversity – convex hull (Mouillot et al. 2021).

From a biological point of view, it is important to remember that the selection of traits should be primarily driven by expert-based considerations of the studied organisms and system(s). From a statistical point of view, approaches for reducing the number of dimensions are no different from those used in other ecological domains (see Dormann et al. 2013 for an overview), and include dropping collinear predictors or reducing the number of correlated traits via sequential regression (= residual regression; Graham 2003) or using dimensionality reduction techniques (e.g. principal coordinate analysis) (Maire et al. 2015).

corrections, e.g. rarefaction of the data or accounting for species' imperfect detection, are needed to remove biases in the data (Box 1); 2) how to handle missing data (Box 2); 3) how to deal with collinearity (e.g. remove collinear traits, reduce dimensionality, identify set of correlated traits to define functional groups) (Box 3); 4) how to handle outliers (in either trait or occurrence data), which might be of interest to the research question (Carmona et al. 2017, Violle et al. 2017) or should be removed being measurement errors; and 5) whether to weight the traits and/or transform them in different ways to comply with the assumptions of the subsequent analyses (see additional insights in Step 6).

The Methods section should include a brief explanation of the problems and decisions made following data exploration.

Step 6. Estimate functional diversity

Only when the sampling design has been set up and implemented (Step 2), the data assembled (Step 3–4), inspected and cleaned (Step 5), it is possible to estimate functional diversity and to evaluate whether statistical relationships exist that can be linked to the primary question of interest (Step 1).

If summarising or comparing univariate trait characteristics is the principal goal of the study, then raw trait data can often be used without any data transformation. The most common examples of a univariate metric that uses raw trait data are the coefficient of variation (standard deviation divided by mean relationship; Yang et al. 2020) and the community-weighted mean (Garnier et al. 2004, Lavorel et al. 2008), which summarises the mean trait value of all individuals or species in the population or assemblage.

If the focus of the study is quantifying multivariate functional diversity, then this is achieved by first constructing a trait space(s) of the study system(s) from the **T** matrix, and then summarising it/them into meaningful descriptive metric(s) after accounting for the information in the **C** matrix. The first step in constructing a trait space is creating a trait dissimilarity matrix for all pairs of individuals or species. Caution must be exercised when choosing a dissimilarity metric, as well as weights for each of the traits. A common practice in trait-based ecology is to assign the same weight to each trait (de Bello et al. 2021, Jarzyna et al. 2021), though consensus on best practices is lacking. For highly dimensional trait data including a combination of continuous, fuzzy coded, categorical and binary traits, the Gower's distance (Gower 1971, Pavoine et al. 2009, de Bello et al. 2021) is the only metric option, because it can handle different types of traits and balances the contribution of traits and trait groups to overall dissimilarity (de Bello et al. 2021).

Several methods exist to construct a trait space from the trait dissimilarity matrix, including functional dendrograms (Petchey and Gaston 2002), convex hulls (Cornwell et al. 2006), and probabilistic hypervolumes (Blonder et al. 2014, Carmona et al. 2016, 2019, Mammola and Cardoso 2020). Functional dendrograms, often created following a clustering procedure which best preserves original distances in the dissimilarity matrix (Mérigot et al. 2010), represent numerical traits fairly accurately, but perform poorly for non-continuous traits and have a strong dependence on the clustering method (Mouchet et al. 2008, Maire et al. 2015). Convex hulls and hypervolumes represent differences based on continuous and non-continuous traits more accurately (Villéger et al. 2008, Laliberte et al. 2010, Blonder et al. 2014), but are computationally more demanding. Regardless of the approach used, when building the trait space it is important to consider possible distortions between initial and final distances. Both dendrogram or hyperspatial representations are subject to possible distortions, and their degree can be tested by checking the correspondence between initial and final distances. This can be achieved, for example, using functions from R packages

'BAT' (*tree.quality* and *hyper.quality*) and 'mFD' (*quality.fspaces*) (Table 1).

Once the trait space is constructed, one can calculate functional diversity metrics suitable to answer research questions at different levels of organisation – individual observations used to construct the trait space, trait space level (alpha FD), pairwise comparisons of trait spaces (beta FD) or the whole system (gamma FD). When calculating multiple components of functional diversity, we advise that researchers are consistent in the construction of the trait space, namely using a single trait space representation for all estimations. A comprehensive characterisation of a trait space typically includes quantifying three components of functional diversity: richness, divergence and regularity (Villéger et al. 2008, Pavoine and Bonsall 2011, Mammola et al. 2021). Functional richness measures the total breadth of functional diversity in a system. For functional dendrograms, functional richness is quantified as a sum of the dendrogram branch lengths (Petchey and Gaston 2006), sometimes weighted by abundance or detection-corrected probability of species occurrence (Jarzyna and Jetz 2016). For convex hulls, functional richness is defined as the volume of the minimum polygon that encloses all species (Mason et al. 2005), and for probabilistic hypervolumes it is a measure of the volume of the hyperspace (Mammola and Cardoso 2020). Functional divergence represents how observations are spread across the occupied trait space (Villéger et al. 2008); it is often quantified as the average distance among observations or the mean distance of species to the centroid of their shared trait space (Villéger et al. 2008, Laliberté and Legendre 2010, Mammola et al. 2021). Lastly, functional regularity reflects the regularity of observations' distribution within the trait space. Among other methods, it can be computed as the regularity of branch lengths in a functional dendrogram (Villéger et al. 2008) or, for hypervolumes, as the overlap between the observed hyperspace and a hypothetical hyperspace where traits and abundances are evenly distributed (Carmona et al. 2016, Mammola and Cardoso 2020). It must be noted that no approach is currently available for estimating divergence and regularity of convex hulls, since a convex hull is a homogeneous (binary) representation of the trait space, by definition equally dispersed and even throughout (see details in Mammola et al. 2021: p. 1873, Table 3).

Note that most approaches to study functional diversity can also integrate intraspecific variation in community-level calculations, including functional dendrograms (Cianciaruso et al. 2009, Cardoso et al. 2015), weighted-abundance sums of trait probability distributions across organisational levels (Carmona et al. 2016, 2019) or the union of functional hypervolumes (Mammola and Cardoso 2020, Graco-Roza et al. 2022) (Step 4).

Step 7. Interpret and validate results

Depending on the primary research question (Step 1), functional diversity metrics (both absolute and those corrected for species richness) might be further used in statistical analyses to link functional diversity with different ecological predictors. A vast number of models are available in the literature, yet most statistical approaches relate functional diversity metrics through space or time to different environmental variables (e.g. generalised additive or linear models, structural equation models, machine learning algorithms, null models). Regardless of the approach, key elements to report include sample size, effect sizes, uncertainty estimates (e.g. standard errors, credible intervals) and model support (e.g. information criteria, variance explained, discriminatory power) (Gerstner et al. 2017). Providing an absolute measure of

Table 1. Examples of R packages and functions (in italics) aiding to implement the eight-step protocol for functional diversity analyses. Note that this list is not exhaustive.

Step	Description	R packages (or functions)
1. Identify an appropriate research question	Literature review, research interest and hypothesis development	litsearchr, redyarn
2. Identify an appropriate study design	Simulations	*simul.comms()*, virtualspecies
3. Assemble a community data matrix	Occurrence data retrieving	rgbif, spocc
	Data manipulation	base, tidyverse
4. Assemble a trait data matrix	Trait data retrieving	BIEN, TR8, rfishbase, arakno
	Data manipulation	dplyr, tidyr, mFD
5. Explore and prepare the data	Data visualisation	base, ggplot2, lattice, plotly, visreg, mFD
	Collinearity	car, usdm, VIF
	Missing data visualisation and imputation	Amelia, BAT, mice, VIM
	Imperfect detection	DiversityOccupancy, unmarked
6. Estimate functional diversity	Data transformation	BAT, FactoMineR, FD, mFD
	Functional diversity metrics computation	adiv, cati, BAT, FD, FDiversity, funrar, hillR, mFD, TPD
7. Interpret and validate the results	Model fit	bmrs, lme4, nlme, glmmTMB, MCMCglmm, mgcv, lavaan, piecewiseSEM, randomForest
	Cross-validation, bootstrapping and jackknifing	CrossValidate, cvTools, bootstrap
	Data visualisation	base, lattice, ggplot2
8. Ensure reproducibility	Cite packages and their version!	base::citation()

model goodness-of-fit is crucial to assess how well it explains or predicts the ecological response(s) (Mac Nally et al. 2018). How to report statistical models is beyond the scope of this paper, and we refer the reader to Zuur and Ieno (2016) for an overview of presenting results in regression-types analyses.

Notably, some descriptors of functional diversity (e.g. functional richness) tend to be closely associated with species richness, such that their interpretation often relies on statistically controlling for this association. Null models are typically built to address this correlation, evaluating whether the observed values of functional diversity metrics deviate from the expectation conditional on the same species richness across sampling units. First, null (i.e. expected) values of each functional diversity metric are obtained by randomly selecting species from a regional species pool and randomising (often > 100 times) species' occurrence values, while keeping site-level species richness constant (Mason et al. 2013). Regional species pools for randomisation can be determined using distance-based clustering analysis (Carstensen et al. 2013) and might differ depending on the type and scale of the analysis. Next, standardised effect sizes (SES) of the deviation of the observed from the expected values are calculated as a measure of significance, with positive and negative values of SES indicating functional diversity higher and lower, respectively, than expected given species richness. Because interpretation of the SES values relies on the assumption of a symmetric null distribution, the skewness of null distributions for each functional diversity metric should be computed (Botta-Dukát 2018). Alternatively, or in addition to SES values, quantile scores and their associated p-values might be quantified for the observed functional diversity values (Swenson 2011). Observed values that fall outside the 2.5% and 97.5% quantiles of the null distribution are indicative of functional diversity higher and lower, respectively, than expected given species richness. For an in-depth discussion on null models, we refer the reader to Götzenberger et al. (2016).

After model fitting, researchers may desire to determine the generality in their results through validation. Validation determines how a model performs across contexts, either through the application to a novel (or partly novel) dataset, or through the comparison of the model's performance with one based on simulations of settings where the process of interest is eliminated, i.e. null models. Validation can help determine the limitations of an analysis in terms of its ability to explain phenomena or to extrapolate to new scenarios, which should then be summarised in the text of the manuscript.

Validation of results in functional diversity analyses should follow standard statistical procedures, which depend on the type of question and model. It is often required to use independent training, validation and testing datasets when the goal is predicting beyond the range of values in the data (e.g. future predictions). Resampling methods such as jackknife or cross-validation are often needed when data are limited or autocorrelated (Roberts et al. 2017), particularly for extrapolation. Because many functional diversity studies take place over large spatial and temporal scales, the validation of models accounting for spatial and temporal autocorrelation is critical (Dormann et al. 2007). For example, quantifying the predictive performance of spatial validation should be an important part in assessing the performance and bias of any modelling results (Ploton et al. 2020).

Step 8. Ensure reproducibility

Proper data curation, management and archival standards should be followed to maximise the transparency and reproducibility of a functional diversity study. The FAIR guiding principles for scientific data management suggest that data should be *Findable*, *Accessible*, *Interoperable* and *Reusable* (Wilkinson et al. 2016). Below, we outline mechanisms that could help the field of trait-based ecology conform to these guiding principles.

Findable data, metadata and code, should be properly documented and referred to by a unique identifier. One way of accomplishing this is through the deposition of data and code used in analyses into an archival/repository service which provides digital object identifiers (DOIs). Static repositories such as Zenodo, Dryad and FigShare are useful for preserving the code used in analysis at the time of publication. Making code *findable* is especially important for functional diversity analyses, given the different ways functional diversity can be calculated (see Step 6). Research is *accessible* through the sharing of these data, metadata and code, typically achieved by linking these to the paper via a Data availability statement. While there are limitations in the types of data that can be shared freely, the use of sample data (i.e. the community and trait matrices used to compute functional diversity) is encouraged within existing data licence agreements (Tulloch et al. 2018). Moreover, whenever possible, open-source protocols should be used ensuring the research is *accessible* in the future, critical for the rapidly growing field of functional diversity.

For data files, fields that contain information should be summarised by metadata that describe the type of data and their origin (Michener et al. 1997). These metadata should be provided with the original, archived data file. This is particularly important for functional diversity, where it is common practice to obtain trait information from different sources. The original sources of data (e.g. those listed in the Supporting information) should be properly referenced and identified allowing for *interoperability* and *reusability* in the future, and database versions, along with download dates, should be specified. An important component of *interoperability* is, wherever possible, adopting standardised practices and vocabulary as they allow for aggregation of heterogeneous sources (Schneider et al. 2019). For example, the Thesaurus of Plant characteristics aims to standardise concepts of plant traits (Garnier et al. 2017). Standardisation of practices, including proper citation of the software and analytical tools used are essential for *interoperability* and *reusability*, ensuring that as increasing sample and trait data become available studies can be reproduced with ease.

Many researchers find themselves thinking about reproducibility after a project is completed – even here, we have

included reproducibility as the final step! – but we stress that FAIR practices should be implemented from a project's inception. The Open Science Framework provides an online platform to link data and code storage systems (including Dropbox, OneDrive, GitHub and their own cloud storage). This architecture allows the merging of hosting platforms more suited for code with more visually-oriented project wiki pages for protocols, methodology and analysis. The use of these stable cloud storage platforms by research groups also ensures long-term availability of all project components within a lab in spite of researchers' turnover.

Web application

To aid researchers in the task of performing trait-based analyses, we developed a Shiny web app that goes through the proposed protocol. The *stepFD* app allows users to check the requirements needed at each step to fully reproduce their study and create a 'reproducibility report' that can be uploaded alongside the published paper to ensure transparent communication of methods and analyses. The checklist can also be downloaded (either as a .csv or .doc file) to be filled out offline. We encourage researchers to start filling out the form before carrying out a study. This will 1) promote thinking of ecological questions and portraying the steps of the project in detail (helpful also when writing grant proposals), 2) reveal potential sampling, data and statistical issues beforehand and 3) implement FAIR practices from the project's inception. The Shiny app, including a user's guide, is available at <https://facuxpalacio.shinyapps.io/stepFD/>.

Conclusions

Our protocol offers a set of simple guidelines aimed at maximising reproducibility, transparency and consistency of functional diversity analyses (Fig. 2). We would like to leave the reader with a few points of reflection.

1) *Be flexible: do not limit yourself.* While the protocol structure may appear dogmatic, our goal is not limiting creativity and lateral thinking. To us, this protocol is a flexible tool to aid researchers in navigating functional diversity analyses and in remembering key pitfalls and steps to transparently document a trait-based study. However, some of the steps presented here may not apply under specific circumstances – e.g. there are cases where it is not advisable to share sensitive data (Tulloch et al. 2018) – and specific research questions may require that one violates some of our recommendations (e.g. night science; Yanai and Lercher 2020).
2) *Be a giant: offer your shoulders.* The correct reporting of methods and statistics, as well as sharing data and codes, provides the foundation for other scientists to build upon your work. A thorough description of sample sizes, statistics and model estimates ensures that others can incorporate your findings into meta-analyses (Gerstner et al. 2017), shared data can be integrated into larger databases for consensus analyses (Mouillot et al. 2021, Graco-Roza et al. 2022), and available code can be reused by other researchers. Whether one sees this altruistically, as a collaborative effort to advance science as a whole, or opportunistically, as a way to increase one' own citations and credibility in the field, the long-term benefits are undisputed.
3) *Be informed: find your way through the jungle of metrics.* As we have shown, functional ecology is a fast-growing field of research (Fig. 1). We have touched upon examples of methods and metrics based on the current literature, but new tools and approaches are being developed continuously, and one must keep up with the literature to make the best out of this field (Mammola et al. 2021). Even though new methods will become available and concepts will emerge in the future, we believe that the key underlying philosophy and motivations of this protocol will remain valid and applicable.
4) *Be permeable: exchange with other disciplines.* Functional diversity represents only one of multiple frameworks within ecology. The constant interaction and integration with other disciplines forming the broader biodiversity research platform (e.g. taxonomy, phylogeny) is fundamental to answer questions and test hypotheses relevant to functional diversity itself.

All in all, we envision our protocol as a set of good practices and starting points; we are convinced that, as other standard protocols did, it may boost effective communication and an enhanced understanding of upcoming functional diversity research.

Acknowledgements – The reproducible code box (Supporting information) was adapted from Joseph R. Bennett's lab manual (Carleton Univ.; compiled by Jaimie G. Vincent), and heavily influenced by the Open Science Foundation.
Funding – FXP received partial support from Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). FR acknowledges support by MITACS through an Accelerate Fellowship (IT23330). GO was supported by the long-term research development project of the Czech Academy of Sciences (RVO 67985939). SM acknowledges support by the European Commission via the Marie Sklodowska-Curie Individual Fellowships program (H2020-MSCA-IF-2019; project no. 882221).
Conflict of interest – The authors have no conflicts of interest to declare.

Author contributions

Facundo Palacio: Conceptualization (lead); Software (lead); Writing – original draft (lead); Writing – review and editing (lead). **Corey T. Callaghan**: Conceptualization (equal); Writing – original draft (equal); Writing – review and editing (equal). **Pedro Cardoso**: Conceptualization (equal); Writing – original draft (equal); Writing – review and editing (equal). **Emma J. Hudgins**: Conceptualization (equal); Software (equal); Writing – original draft (equal); Writing – review and

Glossary

Functional diversity (= trait diversity, FD).

A characterization of life diversity in terms of the diversity of functions (Malaterre et al. 2019). Operationally, any mathematical estimation of the diversity of traits of individuals composing a given group (a community, an ecosystem and so on), from simple measures of trait distributions (means, standard deviation, coefficient of variation, kurtosis) to the plurality of functional diversity indices developed in the last two decades (refer to Mammola et al. 2021 for an overview).

Intraspecific trait variation.

Trait variance of a group of individuals of the same species. It results from phenotypic plasticity or local adaptation of different genotypes along environmental gradients or in response to biotic interactions (e.g. competition or mutualism).

Replicability.

The process of replicating a certain study using different datasets and/or model systems. A lack of replicability occurs when qualitatively different results are obtained applying the same analytical approach.

Reproducibility.

The process of repeating analyses conducted by others. A lack of reproducibility occurs when different results are obtained when re-analysing the data reported in a paper.

Trait.

Any phenotypical entity – morphological, anatomical, ecological, physiological, behavioural, phenological – measured on individual organisms at any scale, from gene to whole organism, and which can be scaled up from individuals to genotype, population, species, community or ecosystem (Violle et al. 2007, Volaire et al. 2020).

Trait resolution.

The coarseness of measured traits, ranging from highest-resolution continuous measurements to lowest-resolution binary categories (Kohli and Jarzyna 2021). Body size measured on a continuous scale is typically a high-resolution trait, whereas the categorical version of this trait (e.g. 'small', 'medium' or 'large') is a low-resolution one.

editing (equal). **Marta A. Jarzyna**: Conceptualization (equal); Writing – original draft (equal); Writing – review and editing (equal). **Gianluigi Ottaviani**: Conceptualization (equal); Writing – original draft (equal); Writing – review and editing (equal). **Federico Riva**: Conceptualization (equal); Writing – original draft (equal); Writing – review and editing (equal). **Caio Graco-Roza**: Conceptualization (equal); Software (equal); Writing – original draft (equal); Writing – review and editing (equal). **Vaughn Shirey**: Conceptualization (equal); Writing – original draft (equal); Writing – review and editing (equal). **Stefano Mammola**: Conceptualization (lead); Data curation (lead); Formal analysis (lead); Visualization (lead); Writing – original draft (equal); Writing – review and editing (equal).

Transparent peer review

The peer review history for this article is available at <https://publons.com/publon/10.1111/ecog.06287>.

Data availability statement

Computer code associated with this publication is available in GitHub, namely the R code and data to generate Fig. 1 (<https://github.com/StefanoMammola/Palacio_et_al_2021_FD_protocol.git>) and the source code for the Shiny app (<https://github.com/facuxpalacio/stepFD>).

Supporting information

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

References

Aglieri, G. et al. 2020. Environmental DNA effectively captures functional diversity of coastal fish communities. – Mol. Ecol. 30: 3127–3139.

Albert, C. H. et al. 2011. When and how should intraspecific variability be considered in trait-based plant ecology? – Perspect. Plant Ecol. Evol. Syst. 13: 217–225.

Andrew, S. C. et al. 2021. Functional diversity of the Australian flora: strong links to species richness and climate. – J. Veg. Sci. 32: e13018.

Araújo, M. B. et al. 2019. Standards for distribution models in biodiversity assessments. – Sci. Adv. 5: eaat4858.

Baker, M. 2016. Is there a reproducibility crisis? – Nature 533: 353–366.

Betts, M. G. et al. 2021. When are hypotheses useful in ecology and evolution? – Ecol. Evol. 11: 5762–5776.
Blasco-Moreno, A. et al. 2019. What does a zero mean? Understanding false, random and structural zeros in ecology. – Methods Ecol. Evol. 10: 949–959.
Blonder, B. et al. 2014. The n-dimensional hypervolume. – Global Ecol. Biogeogr. 23: 595–609.
Bonte, D. et al. 2003. Low propensity for aerial dispersal in specialist spiders from fragmented landscapes. – Proc. R. Soc. B 270: 1601–1607.
Borregaard, M. K. and Hart, E. M. 2016. Towards a more reproducible ecology. – Ecography 39: 349–353.
Botta-Dukát, Z. 2018. Cautionary note on calculating standardized effect size (SES) in randomization test. – Commun. Ecol. 19: 77–83.
Cadotte, M. W. and Tucker, C. M. 2017. Should environmental filtering be abandoned? – Trends Ecol. Evol. 32: 429–437.
Callaghan, C. T. et al. 2021. Three frontiers for the future of biodiversity research using citizen science data. – BioScience 71: 55–63.
Cardoso, P. et al. 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.
Carmona, C. P. et al. 2016. Traits without borders: integrating functional diversity across scales. – Trends Ecol. Evol. 31: 382–394.
Carmona, C. P. et al. 2017. Towards a common toolbox for rarity: a response to Violle et al. – Trends Ecol. Evol. 32: 889–891.
Carmona, C. P. et al. 2019. Trait probability density (TPD): measuring functional diversity across scales based on TPD with R. – Ecology 100: e02876.
Carstensen, D. W. et al. 2013. Introducing the biogeographic species pool. – Ecography 36: 1310–1318.
Christie, A. P. et al. 2019. Simple study designs in ecology produce inaccurate estimates of biodiversity responses. – J. Appl. Ecol. 56: 2742–2754.
Cianciaruso, M. V. et al. 2009. Including intraspecific variability in functional diversity. – Ecology 90: 81–89.
Cordlandwehr, V. et al. 2013. Do plant traits retrieved from a database accurately predict on-site measurements? – J. Ecol. 101: 662–670.
Cornelissen, J. H. C. et al. 2003. A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. – Aust. J. Bot. 51: 335–380.
Cornwell, W. K. et al. 2006. A trait-based test for habitat filtering: convex hull volume. – Ecology 87: 1465–1471.
Culina, A. et al. 2020. Low availability of code in ecology: a call for urgent action. – PLoS Biol. 18: e3000763.
Currie, D. J. 2019. Where Newton might have taken ecology. – Global Ecol. Biogeogr. 28: 18–27.
Dawson, S. K. et al. 2021. The traits of 'trait ecologists': an analysis of the use of trait and functional trait terminology. – Ecol. Evol. 11: 16434–16445.
de Bello, F. et al. 2012. Functional species pool framework to test for biotic effects on community assembly. – Ecology 93: 2263–2273.
de Bello, F. et al. 2021. Towards a more balanced combination of multiple traits when computing functional differences between species. – Methods Ecol. Evol. 12: 443–448.
de Jager, M. et al. 2019. Seed size regulates plant dispersal distances in flowing water. – J. Ecol. 107: 307–317.
Denes, F. V. et al. 2015. Estimating abundance of unmarked animal populations: accounting for imperfect detection and other sources of zero inflation. – Methods Ecol. Evol. 6: 543–556.
Denny, M. 2017. The fallacy of the average: on the ubiquity, utility and continuing novelty of Jensen's inequality. – J. Exp. Biol. 220: 139–146.
Devarajan, K. et al. 2020. Multi-species occupancy models: review, roadmap and recommendations. – Ecography 43: 1612–1624.
Díaz, S. et al. 2016. The global spectrum of plant form and function. – Nature 529: 167–171.
Dormann, C. F. et al. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. – Ecography 30: 609–628.
Dormann, C. F. et al. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. – Ecography 36: 27–46.
Eckert, E. M. et al. 2020. Every fifth published metagenome is not available to science. – PLoS Biol. 18: e3000698.
Ellison, A. M. 2010. Repeatability and transparency in ecological research. – Ecology 91: 2536–2539.
Fanelli, D. 2018. Opinion: Is science really facing a reproducibility crisis, and do we need it to? – Proc. Natl Acad. Sci. USA 115: 2628–2631.
Feng, X. et al. 2019. A checklist for maximizing reproducibility of ecological niche models. – Nat. Ecol. Evol. 3: 1382–1395.
Fidler, F. et al. 2017. Metaresearch for evaluating reproducibility in ecology and evolution. – BioScience 67: 282–289.
Fraser, H. et al. 2018. Questionable research practices in ecology and evolution. – PLoS One 13: e0200303.
Garnier, E. et al. 2004. Plant functional markers capture ecosystem properties during secondary succession. – Ecology 85: 2630–2637.
Garnier, E. et al. 2017. Towards a thesaurus of plant characteristics: an ecological contribution. – J. Ecol. 105: 298–309.
Gasc, A. et al. 2013. Assessing biodiversity with sound: do acoustic diversity indices reflect phylogenetic and functional diversities of bird communities? – Ecol. Indic. 25: 279–287.
Gentile, G. et al. 2021. Evaluating intraspecific variation in insect trait analysis. – Ecol. Entomol. 46: 11–18.
Gerstner, K. et al. 2017. Will your paper be used in a meta-analysis? Make the reach of your research broader and longer lasting. – Methods Ecol. Evol. 8: 777–784.
Gomez, J. P. et al. 2018. An efficient extension of N-mixture models for multi-species abundance estimation. – Methods Ecol. Evol. 9: 340–353.
Götzenberger, L. et al. 2016. Which randomizations detect convergence and divergence in trait-based community assembly? A test of commonly used null models. – J. Veg. Sci. 27: 1275–1287.
Gower, J. C. 1971. A general coefficient of similarity and some of its properties. – Biometrics 27: 857–871.
Graco-Roza, C. et al. 2022. Distance decay 2.0 – a global synthesis of taxonomic and functional turnover in ecological communities. – Global Ecol. Biogeogr. 31: 1399–1421.
Graham, M. H. 2003. Confronting multicollinearity in ecological multiple regression. – Ecology 84: 2809–2815.
Greenop, A. et al. 2018. Functional diversity positively affects prey suppression by invertebrate predators: a meta-analysis. – Ecology 99: 1771–1782.
Günter, F. et al. 2019. Latitudinal and altitudinal variation in ecologically important traits in a widespread butterfly. – Biol. J. Linn. Soc. 128: 742–755.

Hastings, A. et al. 2005. The spatial spread of invasions: new developments in theory and evidence. – Ecol. Lett. 8: 91–101.
Hurlbert, A. H. and Jetz, W. 2007. Species richness, hotspots and the scale dependence of range maps in ecology and conservation. – Proc. Natl Acad. Sci. USA 104: 13384–13389.
Iknayan, K. J. et al. 2014. Detecting diversity: emerging methods to estimate species diversity. – Trends Ecol. Evol. 29: 97–106.
Jarzyna, M. A. and Jetz, W. 2016. Detecting the multiple facets of biodiversity. – Trends Ecol. Evol. 31: 527–538.
Jarzyna, M. A. et al. 2021. Global functional and phylogenetic structure of avian assemblages across elevation and latitude. – Ecol. Lett. 24: 196–207.
Johnson, T. F. et al. 2021. Handling missing values in trait data. – Global Ecol. Biogeogr. 30: 51–62.
Klimešová, J. et al. 2019. Handbook of standardized protocols for collecting plant modularity traits. – Perspect. Plant Ecol. Evol. Syst. 40: 125485.
Kohli, B. A. and Jarzyna, M. A. 2021. Pitfalls of ignoring trait resolution when drawing conclusions about ecological processes. – Global Ecol. Biogeogr. 30: 1139–1152.
Laliberté, E. and Legendre, P. 2010. A distance-based framework for measuring functional diversity from multiple traits. – Ecology 91: 299–305.
Lancaster, J. and Downes, B. J. 2017. Dispersal traits may reflect dispersal distances, but dispersers may not connect populations demographically. – Oecologia 184: 171–182.
Landhuis, E. 2016. Scientific literature: information overload. – Nature 535: 457–458.
Lavorel, S. et al. 2008. Assessing functional diversity in the field–methodology matters! – Funct. Ecol. 22: 134–147.
Luck, G. W. et al. 2012. Improving the application of vertebrate trait-based frameworks to the study of ecosystem services. – J. Anim. Ecol. 81: 1065–1076.
Mac Nally, R. et al. 2018. Model selection using information criteria, but is the 'best' model any good? – J. Appl. Ecol. 55: 1441–1444.
Maire, E. et al. 2015. How many dimensions are needed to accurately assess functional diversity? A pragmatic approach for assessing the quality of functional spaces. – Global Ecol. Biogeogr. 24: 728–740.
Malaterre, C. et al. 2019. Functional diversity: an epistemic roadmap. – BioScience 69: 800–811.
Mammola, S. and Cardoso, P. 2020. Functional diversity metrics using kernel density n-dimensional hypervolumes. – Methods Ecol. Evol. 11: 986–995.
Mammola, S. et al. 2021. Concepts and applications in functional diversity. – Funct. Ecol. 35: 1869–1885.
Mason, N. W. et al. 2005. Functional richness, functional evenness and functional divergence: the primary components of functional diversity. – Oikos 111: 112–118.
Mason, N.W.H. et al. 2013. A guide for using functional diversity indices to reveal changes in assembly processes along ecological gradients. – J. Veg. Sci. 24: 794–806.
Matuoka, M. A. et al. 2020. Effects of anthropogenic disturbances on bird functional diversity: a global meta-analysis. – Ecol. Indic. 116: 106471.
McGill, B. J. 2010. Matters of scale. – Science 328: 575–576.
McGill, B. J. et al. 2006. Rebuilding community ecology from functional traits. – Trends Ecol. Evol. 21: 178–185.
McPherson, J. M. et al. 2018. A simulation tool to scrutinise the behaviour of functional diversity metrics. – Methods Ecol. Evol. 9: 200–206.
Mentis, M. T. 1988. Hypothetico-deductive and inductive approaches in ecology. – Funct. Ecol. 2: 5–14.
Mérigot, B. et al. 2010. On goodness-of-fit measure for dendrogram-based analyses. – Ecology 91: 1850–1859.
Michener, W. K. et al. 1997. Nongeospatial metadata for the ecological sciences. – Ecol. Appl. 7: 330–342.
Moretti, M. et al. 2017. Handbook of protocols for standardized measurement of terrestrial invertebrate functional traits. – Funct. Ecol. 31: 558–567.
Mouchet, M. et al. 2008. Towards a consensus for calculating dendrogram-based functional diversity indices. – Oikos 117: 794–800.
Mouillot, D. et al. 2021. The dimensionality and structure of species trait spaces. – Ecol. Lett. 24: 1988–2009.
Munafò, M. R. et al. 2017. A manifesto for reproducible science. – Nat. Hum. Behav. 1: 21.
Nakagawa, S. and Freckleton, R. 2008. Missing inaction: the dangers of ignoring missing data. – Trends Ecol. Evol. 23: 592–596.
Newbold, T. et al. 2020. Global effects of land use on biodiversity differ among functional groups. – Funct. Ecol. 34: 684–693.
Palacio, F. X. et al. 2019a. Does accounting for within-individual trait variation matter for measuring functional diversity? – Ecol. Indic. 102: 43–50.
Palacio, F. X. et al. 2019b. Measuring natural selection on multivariate phenotypic traits: a protocol for verifiable and reproducible analyses of natural selection. – Israel J. Ecol. Evol. 65: 130–136.
Palacio, F. X. et al. 2020. The costs of ignoring species detectability on functional diversity estimation. – Auk 137: ukaa057.
Pausas, J. G. et al. 2016. Towards understanding resprouting at the global scale. – New Phytol. 209: 945–954.
Pavoine, S. and Bonsall, M. B. 2011. Measuring biodiversity to explain community assembly: a unified approach. – Biol. Rev. 86: 792–812.
Pavoine, S. et al. 2009. On the challenge of treating various types of variables: application for improving the measurement of functional diversity. – Oikos 118: 391–402.
Penone, C. et al. 2014. Imputation of missing data in life-history trait datasets: which approach performs the best? – Methods Ecol. Evol. 5: 961–970.
Perez, T. M. et al. 2020. Herbarium-based measurements reliably estimate three functional traits. – Am. J. Bot. 107: 1457–1464.
Perronne, R. et al. 2017. How to design trait-based analyses of community assembly mechanisms: insights and guidelines from a literature review. – Perspect. Plant Ecol. Evol. Syst. 25: 29–44.
Petchey, O. L. and Gaston, K. J. 2006. Functional diversity: back to basics and looking forward. – Ecol. Lett. 9: 741–758.
Petchey, O.L. and Gaston, K.J. 2002. Functional diversity (FD), species richness and community composition. – Ecol. Lett. 5: 402–411.
Pigot, A. L. et al. 2020. Macroevolutionary convergence connects morphological form to ecological function in birds. – Nat. Ecol. Evol. 4: 230–239.
Platt, J. R. 1964. Strong inference. – Science 146: 347–353.
Ploton, P. et al. 2020. Spatial validation reveals poor predictive performance of large-scale ecological mapping models. – Nat. Commun. 11: 1–11.
Podani, J. et al. 2021. Principal component analysis of incomplete data – a simple solution to an old problem. – Ecol. Inform. 61: 101235.

Powers, S. M. and Hampton, S. E. 2019. Open science, reproducibility and transparency in ecology. – Ecol. Appl. 29: e01822.
Ricotta, C. et al. 2012. Functional rarefaction for species abundance data. – Methods Ecol. Evol. 3: 519–525.
Riva, F. et al. 2020. Composite effects of cutlines and wildfire result in fire refuges for plants and butterflies in boreal treed peatlands. – Ecosystems 23: 485–497.
Roberts, D.R. et al. 2017. Cross-validation strategies for data with temporal, spatial, hierarchical or phylogenetic structure. – Ecography 40: 913–929.
Roswell, M. et al. 2021. A conceptual guide to measuring species diversity. – Oikos 130: 321–338.
Roth, T. et al. 2018. Functional ecology and imperfect detection of species. – Methods Ecol. Evol. 9: 917–928.
Scheiner, S. M. and Gurevitch, J. 2001. Design and analysis of ecological experiments. – Oxford Univ. Press.
Schleuter, D. et al. 2010. A user's guide to functional diversity indices. – Ecol. Monogr. 80: 469–484.
Schneider, F. D. et al. 2017. Mapping functional diversity from remotely sensed morphological and physiological forest traits. – Nat. Commun. 8: 1–12.
Schneider, F. D. et al. 2019. Towards an ecological trait-data standard. – Methods Ecol. Evol. 10: 2006–2019.
Shirey, V. et al. 2022. Utilizing occupancy-detection models with museum specimen data: promise and pitfalls. – Methods Ecol. Evol., <https://doi.org/10.1111/2041-210X.13896>.
Sigsgaard, E. E. et al. 2020. Environmental DNA metabarcoding of cow dung reveals taxonomic and functional diversity of invertebrate assemblages. – Mol. Ecol. 30: 3374–3389.
Sobral, M. 2021. All traits are functional: an evolutionary viewpoint. – Trends Plant Sci. 7: 674–676.
Spasojevic, M. J. and Suding, K. N. 2012. Inferring community assembly mechanisms from functional diversity patterns: the importance of multiple assembly processes. – J. Ecol. 100: 652–661.
Swenson, N. G. 2011. Phylogenetic beta diversity metrics, trait evolution and inferring the functional beta diversity of communities. – PLoS One 6: e21264.
Taugourdeau, S. et al. 2014. Filling the gap in functional trait databases: use of ecological hypotheses to replace missing data. – Ecol. Evol. 4: 944–958.
Tenopir, C. et al. 2011. Data sharing by scientists: practices and perceptions. – PLoS One 6: e21101.
Tosa, M. I. et al. 2021. The rapid rise of next-generation natural history. – Front. Ecol. Evol. 9: 698131.
Tulloch, A. I. et al. 2018. A decision tree for assessing the risks and benefits of publishing biodiversity data. – Nat. Ecol. Evol. 2: 1209–1217.
Villéger, S. et al. 2008. New multidimensional functional diversity indices for a multifaceted framework in functional ecology. – Ecology 89: 2290–2301.
Violle, C. et al. 2007. Let the concept of trait be functional! – Oikos 116: 882–892.
Violle, C. et al. 2012. The return of the variance: intraspecific variability in community ecology. – Trends Ecol. Evol. 27: 244–252.
Violle, C. et al. 2014. The emergence and promise of functional biogeography. – Proc. Natl Acad. Sci. USA 111: 13690–13696.
Violle, C. et al. 2017. Functional rarity: the ecology of outliers. – Trends Ecol. Evol. 32: 356–367.
Volaire, F. et al. 2020. What do you mean 'functional' in ecology? Patterns versus processes. – Ecol. Evol. 10: 11875–11885.
Walker, S. C. et al. 2008. Functional rarefaction: estimating functional diversity from field data. – Oikos 117: 286–296.
Weiher, E. et al. 1999. Challenging Theophrastus: a common core list of plant traits for functional ecology. – J. Veg. Sci. 10: 609–620.
Wilkinson, M. D. et al. 2016. The FAIR guiding principles for scientific data management and stewardship. – Sci. Data 3: 1–9.
Wong, M. K. and Carmona, C. P. 2021. Including intraspecific trait variability to avoid distortion of functional diversity and ecological inference: lessons from natural assemblages. – Methods Ecol. Evol. 12: 946–957.
Woodcock, B. A. et al. 2019. Meta-analysis reveals that pollinator functional diversity and abundance enhance crop pollination and yield. – Nat. Commun. 10: 1–10.
Wulff, J. N. and Jeppesen, L. E. 2017. Multiple imputation by chained equations in praxis: guidelines and review. – Elect. J. Busin. Res. Methods 15: 41–56.
Yanai, I. and Lercher, M. 2020. A hypothesis is a liability. – Gen. Biol. 21: 231.
Yang, J. et al. 2020. Large underestimation of intraspecific trait variation and its improvements. – Front. Plant Sci. 11: 53.
Zurell, D. et al. 2020. A standard protocol for reporting species distribution models. – Ecography 43: 1261–1277.
Zuur, A. F. and Ieno, E. N. 2016. A protocol for conducting and presenting results of regression-type analyses. – Methods Ecol. Evol. 7: 636–645.
Zuur, A. F. et al. 2010. A protocol for data exploration to avoid common statistical problems. – Methods Ecol. Evol. 1: 3–14.