Target journal: Functional ecology

**Title**

**Matching evolutionary distant organisms: a systematic review on the use of cross-taxa perspective in functional ecology**

André Luís Luza1, Cesar Cordeiro, Diego Barneche, Juan Pablo Quimbayo2, Fernanda Silva, Vinicius Giglio, Thiago Mendes, Luiza Waechter1, Vitor Picolotto, Cristian Dambros1, Osmar Luiz, Ronaldo Francini-Filho2, Mariana Gomes Bender1, Guilherme Longo

1 - Marine Macroecology and Conservation Lab. Universidade Federal de Santa Maria. Santa Maria - RS. Brazil.

2 - CEBIMar, USP. São Sebastião - SP. Brazil.

3 - Marine Ecology Lab. Universidade Federal do Rio Grande do Norte. Natal - RN. Brazil.

4 – UFSC

5 – AIMS

Authors’ contribution:

ALL executed the literature survey, organized the framework for filtering, extracting, and analyzing data from the surveyed papers.

ALL, CC, DB, FS, JPQ, VG, TM, LW, VP, MB, GOL fully screened the selected papers, and collected information from them.

ALL reviewed collected data, wrote the first draft, and all authors had a critical contribution to improve it through writing and analysis suggestion.

Funding:

This study was granted by CNPq. This study is part of the project ‘Brazilian reef systems in the Anthropocene’, from the Brazilian Institute *SINBIOSE*, The Brazilian center for Synthesis in Biodiversity and Ecosystem Services.

André Luza: Process number 152410/2020-1.

**Introduction**

Looking nature from the perspective of functional traits has reinvented ecology. The emergence of functional ecology has changed the focus from which species to which ecological process (Dias & Cabido 2001, McGill et al. 2006, Violle et al. 2007, 2012, Diaz et al. 2016; Bellwood et al. 2019). Functional ecology has generally contributed to the understanding of processes underlying species distribution (Pavoine & Bonsall 2011, Pausas & Verdú 2010) and diversification (Floeter et al. 2018, Siqueira et al. 2020), the strength of species coupling into food webs (Bugoni et al. 2020), and the contribution of individuals and species to community structure (Violle et al. 2007, 2012, Diaz et al. 2016, Cooke et al. 2019, Pimiento et al. 2020, Cianciaruso et al. 2009). Beyond these topics, functional ecology has revealed the biology behind extinction threat on ecological networks (Bastazini et al. 2018) and ecological communities (Cooke et al. 2019, 2020, Pimiento et al. 2020, Brodie et al. 2021, Carmona et al. 2021, Cerreta et al. 2020), with great potential to be used in conservation (Miatta et al. 2021) and restoration proposes (Carlucci et al. 2020). Although functional ecology has clearly brought new perspectives and questions to ecology (McGill et al. 2006, Mouillot et al. 2013), a major difficulty still remains: to effectively cross the boundaries imposed by taxonomy (Aubin et al. 2013). Although the field of functional ecology is intended to use a ‘taxon-free approach’ (e.g., Lavorel et al. 2008), questions such as ‘how frequently research in functional ecology crosses taxonomic barriers’ and ‘what types of approaches and traits are used when these barriers are overcome’ remain virtually unanswered.

Taking a step forward to fill this gap, Weiss & Ray (2019) recently brought to the ecological literature the concept of ‘functionally analogous traits’, which is the essence of a cross-taxa perspective. Functional analogues are traits that capture similar biological phenomena across multiple taxa (Weiss & Ray 2019). While these traits can be represented by structures of similar and/or different ontogenetic origin across organisms, they must represent a similar response to the environment and/or function in the ecosystem (Weiss & Ray 2019). For instance, in a pioneer study of cross-taxa research in functional ecology, Moretti and Legg (2009) combined traits of plants and animals to show parallel strategies across taxa regarding survival, recovery and dispersal front to disturbances. The emphasis on a cross-taxa perspective is therefore on what organisms do in the ecosystem regardless of how they do and who does it (Bellwood et al. 2019; Weiss & Ray 2019; Tavares et al. 2019). Weiss & Ray (2019) also synthesized a series of advancements in finding, selecting, collating, and comparing response traits -- those influencing organism performance (Diaz & Cabido 2001, Miatta et al. 2021) -- across taxa, as well as presented caveats and future directions in this research agenda. However, how often this approach is used in functional ecology need to be better elucidated.

Here we reviewed the ecological literature with the objective of finding: i) the major taxonomic names and ranks being studied, ii) which links this research field makes in the tree of life, iii) the predominant research context, analytical procedures and functional approaches used in the field, iv) if there is a trade-off between the number of taxa and the number of traits per study in cross-taxa research in functional ecology. Finally, by knowing the ‘state-of-art’ of cross-taxa research in functional ecology, we extend the framework of Mouillot et al. (2013) — regarding single-taxon variation in fitness with disturbances — to embrace multiple taxa, which will let us closer to understand emergent properties of whole ecosystems.

Increasing N of studies (Brosseau et al. 2017).

Open data ((Brosseau et al. 2017), weiss & ray

Generalization across taxa

**Methods**

We searched for literature of cross-taxa research in functional ecology using the main collection of studies from the Web of Science (Clarivate Analytics 2020). The keywords “ecolog\*” and “funct\*” and “trait\*” were used to find research in ecology and functional ecology, together with the keywords “cross-tax\*” or “multi-tax\*” or “multiple-tax\*” or “taxo\*-free” that were used to find research with more than one taxon (Fig. 1). The literature survey returned 1,006 studies published from 1945 to the survey day (2020-06-08). The title, key words and abstract were screened in order to find whether studies covered multiple taxa, and covered topics in functional ecology. Several exclusion criteria were applied in this initial phase of screening (Table S1), including the exclusion of duplicated records (five articles). Details on the criteria applied in this first filtering phase can be found in Supporting Information Table S1. The data supporting this study was collected according to the PRISMA protocol for reporting procedures of systematic reviews (Liberati et al. 2009).

To be eligible to data extraction an article should be:

1. Ecological research: Research explicitly interested in the relationship between organisms and their environment (Ricklefs 1980). Articles on paleoecology, phylogeography and biogeography, with explicit test of ecological processes underlying diversity, distribution, and evolution, also were considered as ecological research (Table S1).
2. Functional ecology research: study using traits and / or functional groups to classify species and assess their response to and/or effect on the environment (Arnold 1983, Violle et al. 2007, Lavorel et al. 2008, refs). We were inclusive here by considering studies using functional groups, because attributing one species to a functional group (e.g., insectivore guild) demands knowledge on a set of traits of the organisms being classified (e.g., feeding behavior, diet content).
3. Cross taxa research: research including community and trait data of multiple taxa from at least one different taxonomic family. We used family because organisms tend to be ecologically and evolutionarily independent units from this to superior taxonomic ranks (Wiens & Graham 2005, Hadley et al. 2009, Anderegg et al. 2018).

We collected data on 26 descriptors of each article that passed by the initial filtering phase (Table S2). With these descriptors we built a data set where each entry consists on descriptors (columns) per trait, taxon and article (rows).

**Data analysis**

*i) The taxonomic names and ranks being studied:* data were aggregated at level of article and taxon, which enabled to find the number of articles per taxon. We analyzed the number of taxonomic ranks by aggregating data at the level of rank instead of taxon. These results were represented through barplots built using the ‘ggplot2’ package of R. All analyzes were ran in R v.4.1.1 (R Core Team 2021).

We analyzed the most frequent links between taxa by building an interaction graph. We used the complete dataset to find all possible combinations of taxa, and then represented only those combinations with one of more articles. Each link refers to the connection between pairs of taxa, and the strength of the link was given by the number articles per pair of taxa. These analyzes were run using functions of the ‘igraph’ package of R.

*ii) What links this research field promotes in the tree of life:* We used the classification of all organisms on Earth of Ruggiero et al. (2015), at the taxonomic level of class, to identify the whole set of possible taxa that could be used in the cross-taxa research. We gathered data on all taxonomic ranks up to the level of class, based on the National Center for Biotechnology Information (NCBI) data base, using the function *classification* of the ‘taxize’ package of R. Then, we bind on it the taxonomic ranks of taxa in our data set (at the level of class, domain and kingdom, the most frequent ranks), and built a topological tree of life based on the similarities of taxonomic ranks across taxa. These similarities were based on Jaccard distance between taxa, and the topology based on a clustering algorithm of UPGMA (Legendre & Legendre 1998). Studied taxa were mapped in the topology by exploring the identity of all edges (tips and internal nodes) of the tree and marking the edges belonging to the taxonomic path of each taxon, from the tip to the root of the topology. This was done by building an adjacency matrix with edges in the columns and tips in the rows; 1s depicted the edges belonging to the taxonomic path of each taxon. Jaccard distance and UPGMA were calculated using functions of the ‘vegan’ package, the topology was handled using functions of the ‘phytools’ package, and the adjacency matrix was built using functions of the ‘daee’ package of R.

*iii) The predominant research context, analytical procedures and functional approaches used in the field:*

The main research contexts and analytical procedures used in this research field were explored through word clouds. Research context refers to the major objective and application of the research, and was based on a broad list of possible contexts of research in ecology (Supporting information). Analytical procedures refer to specific indices and analyses used in the studies. Since one unique article can fit to different research contexts and have used several indices and analysis at once, the word cloud represents the total number of times that a given research context and analytical procedure appeared in the whole dataset (at the level of data entry (trait per study)).

To show how often the authors use the approach of functionally analogous traits (Weiss & Ray 2019), use phylogenies, and use abundance in their articles, we aggregated the data set at the level of article title. Regarding abundance, we further explored, for those indices that can be weighted by abundance (e.g., FEve, FDiv, FDis, Rao’s Quadratic Entropy, Community Weighted Means) (Villéger et al. 2008, Mouquet et al. 2010), the frequency that analyses involved abundance, as the strength of a function is linked to abundance (Mouillot et al. 2013).

To find the predominant functional approaches used in this field of research, we analyzed whether the articles used combine and/or compare data across taxa. The approach of combining data is most common when using functionally analogous traits are used whereby data from different taxa are easily collated and analyzed in a unified way (e.g., Moretti & Legg 2009, Weiss & Ray 2019). Collating data is also a common procedure in ecomorphospace analysis (e.g., ref). In turn, the comparative approach is more common when several taxa are analyzed in parallel and compared in the last analytical step. This kind of approach is common in analysis of ecological congruence and surrogate taxa (e.g., Heino 2010). Finally, studies using both approaches are also possible (hereafter ‘both’) as researchers can be interested in e.g., ecological congruence and emergent properties of the ecosystem (e.g.,).

We analyzed this classification of functional approaches in parallel with definitions of trait categories, in order to understand which type of trait is used in different functional approaches. As there is a broad diversity of traits used in the surveyed research, we categorized traits in 12 categories: diet, habitat, size, reproduction, dispersal, behavior, habit, growth, defense, distribution, interactions, other (see Table S3 to see the traits included in each category). We used a Sankey plot to represent the relationship between approaches and trait categories. In this plot, the width of the connector represents the number of data entries per combination of functional approach and trait category, standardized by the number of data entries per functional approach. The Sankey plot was built using functions of ‘network3D’ package of R.

*iv) Identify a trade-off between the number of taxa and the number of traits per study.*

The number of trait categories and the number of taxa, both aggregated at the level of article title, were used to fit a Poisson model to detect a trade-off in the relationship between number of taxa and traits. We tested three models, each one with first, second, and third order polynomials of the number of taxa, and used model selection analysis based on Akaike Information Criterion to find the model best fitted to the data. We ran the same analysis, for articles with at least two taxa, replacing the number of taxa by the mean pairwise phylogenetic distance between taxa (MPD, Webb et al. 2002). These models were built using functions of the ‘stats’ package of R. Model selection analysis was done with functions implemented in the ‘MuMIn’ package of R.

**Results**

The systematic review resulted in a dataset with 1356 entries, comprising 138 taxa and 496 traits from 97 different articles (~10% of the total number of 1,006 records) (Fig. 1). We found that most cross-taxa research used data of organisms from different classes, phyla and orders (Fig. 2A), and generally used data from one to three taxonomic ranks (Fig. 2B). Tracheophyta, Aves, Eutheria, Coleoptera, and Araneae were the most often cited taxonomic ranks in the ***surveyed*** articles (Fig. 2C). The taxa most often linked in cross-taxa research were Aves, Mammalia, Araneae, Streptophytina (class of plants), Neoptera (wing insects), and Fungi (Fig. 2D). Interestingly, we found that a group of taxa composed by dinoflagellates, cyanobacteria, and other microorganisms did not appear in the same research (did not link) with e.g., higher plants, vertebrates and invertebrates (Fig. 2D). The topologic tree of this research field shows that most research concentrates in several groups of vertebrates and invertebrates, although this can be an artifact of the smaller clade size of these organisms relative to plants, bacteria and other microorganisms (Fig. 3).

In terms of research context, most research is concerned with ‘community structure’, ‘ecosystem monitoring’, and ‘ecological monitoring’ (Fig. 4A). Yet functional ecology is field of research dedicated to the development of indices and frameworks to organisms’ response to and effect on their environment (Petchey & Gaston 2006, Villéger et al. 2008, Mouillot et al. 2013, Mammola & Cardoso 2021), the surveyed research largely relies on taxonomic indexes such as ‘species richness’, ‘abundance’ and ‘taxonomic indices’ (e.g., Shannon, Simpson, Margalef) to count the number of species, individuals and diversity within discretely defined functional groups. Functional composition (e.g., community-weighted means) and functional dispersion (including FDis and Rao’s Entropy) stand out among the most used functional indices (Fig. 4A).

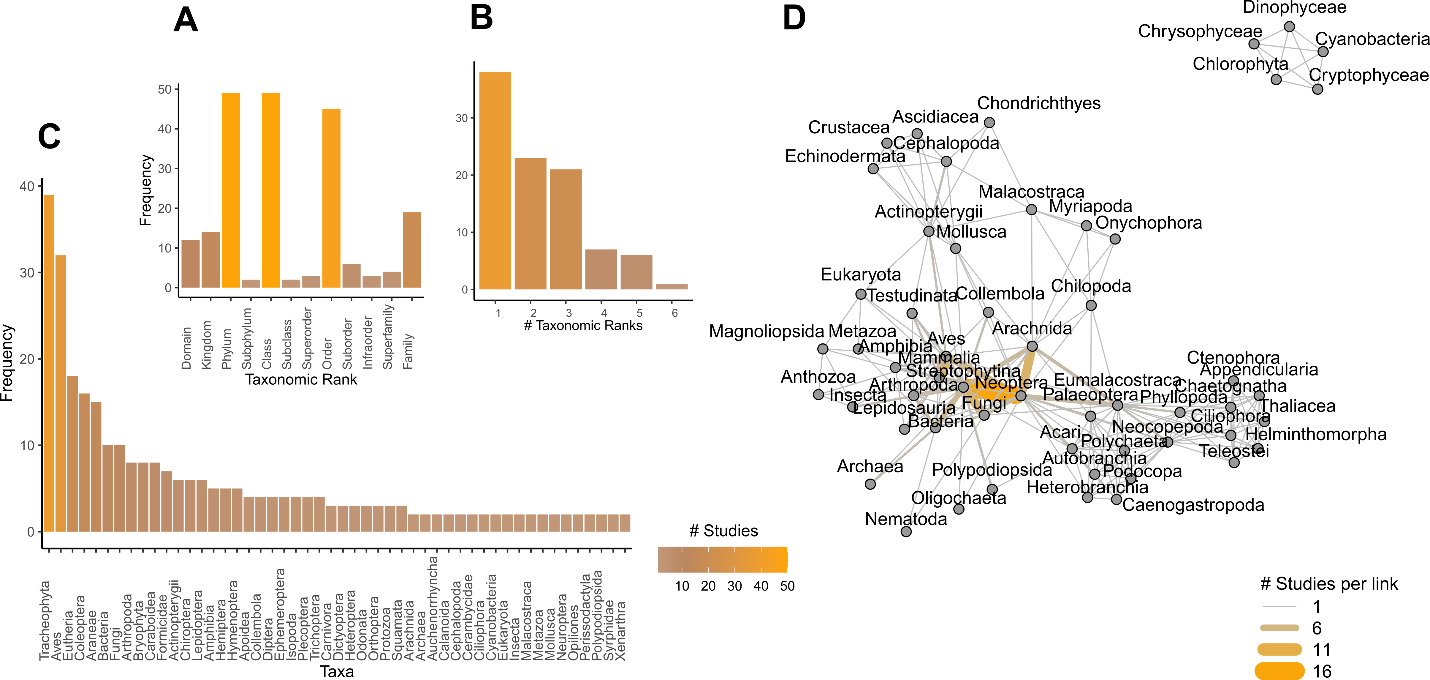
Functionally analogous traits are very often used in surveyed research (Fig. 4B) meaning that authors are striving to represent similar processes across taxa, yet most research still compare data (Fig. 4C). Phylogenies are rarely used in cross-taxa research (Fig. 4B). Phylogeny can be an important aspect to be controlled in analyses, since phylogenetic non-independence between organisms can be present whereby closely related taxa are expected to have more similar responses to environmental gradients than the distantly related ones (Pagel et al. 1997). Also, a few studies explicitly analyze phylogenetic signal in traits (e.g., which used the approach of Blomberg et al. 2003) or yet the phylogenetic structure of communities (e.g., which used the approach described in Webb et al. 2002). Use of abundance was also frequent in the surveyed articles, with half of them using abundance (Fig. 4B). However, by checking which studies used indexes such as functional evenness, functional divergence, functional dispersion, and community weighted means (which can be weighted by abundance), we found only 20 such studies, and 15 of them weighted indices by abundance.

Yet often using functionally analogous traits, most surveyed research compares rather than combines data across taxa (Fig. 4C). This finding means that authors are not fully exploring the approach of functionally analogous traits (Weiss & Ray 2019). In fact, the framework of Weiss & Ray is very recent, so that authors were using trait data for representing similar processes across taxa without a theoretical framework for combining data of different taxa. The percentage of data entries relative to entries per functional approach shows that research combining data often used diet (25.8%), size (19.64%), habitat (16.54%), growth (15.5%) and distribution (4.90%) as trait categories, whereas the research comparing data often used size (21.12%), diet (18.65%), habitat (18.48%), growth (11.55%) and reproduction (10.23%) as trait categories in functional analyses (Fig. 4C, Table S3).

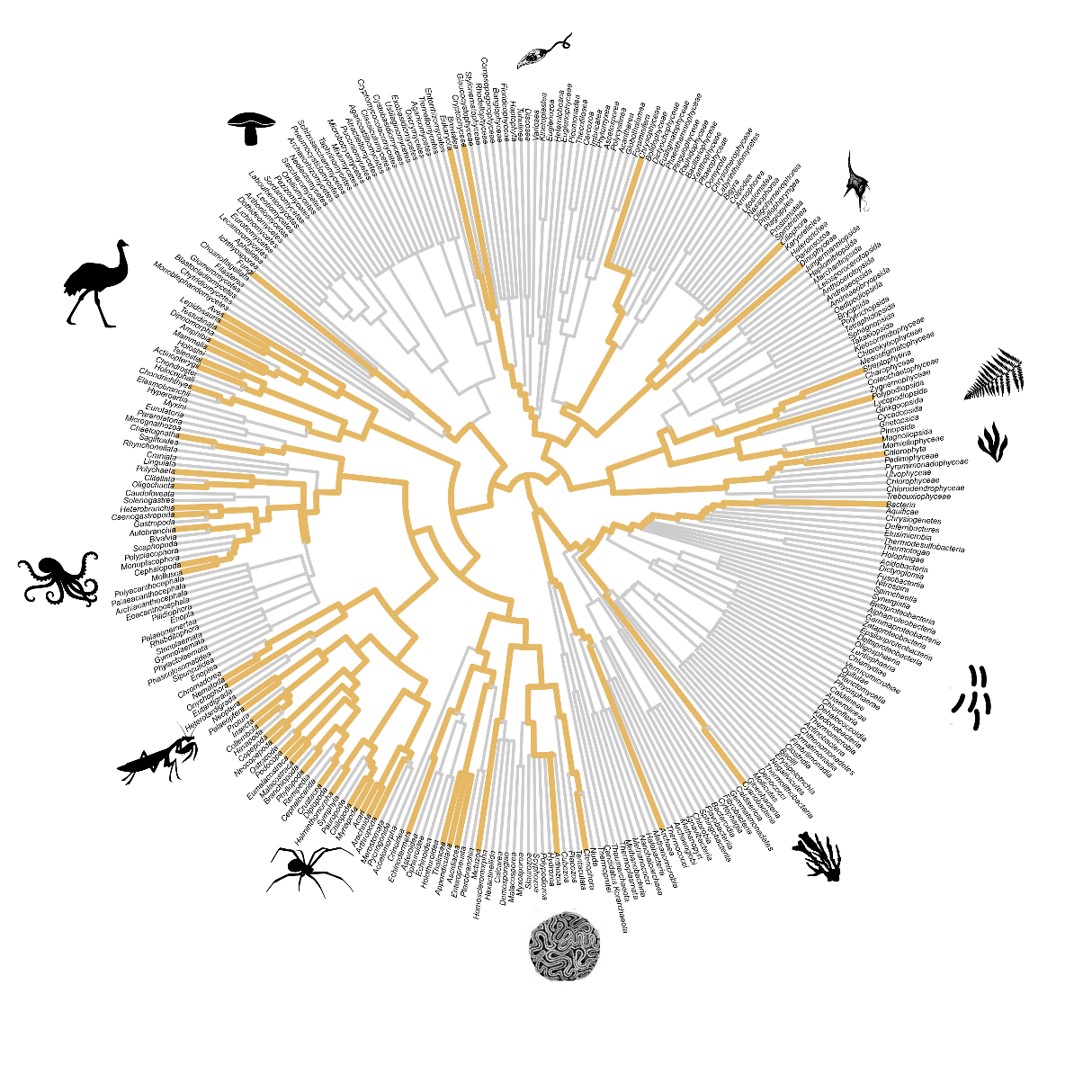
Finally, we found that the models with third and second order polynomials of the number of taxa were the most supported by the data (Table S4), both showing a peak on number of trait categories at five taxa (Figs 5A and S1). The second order polynomial of MPD was the model most supported by the data, showing a peak on the number of trait categories at ~0.6 taxonomic distance between taxa, Fig. 5B).



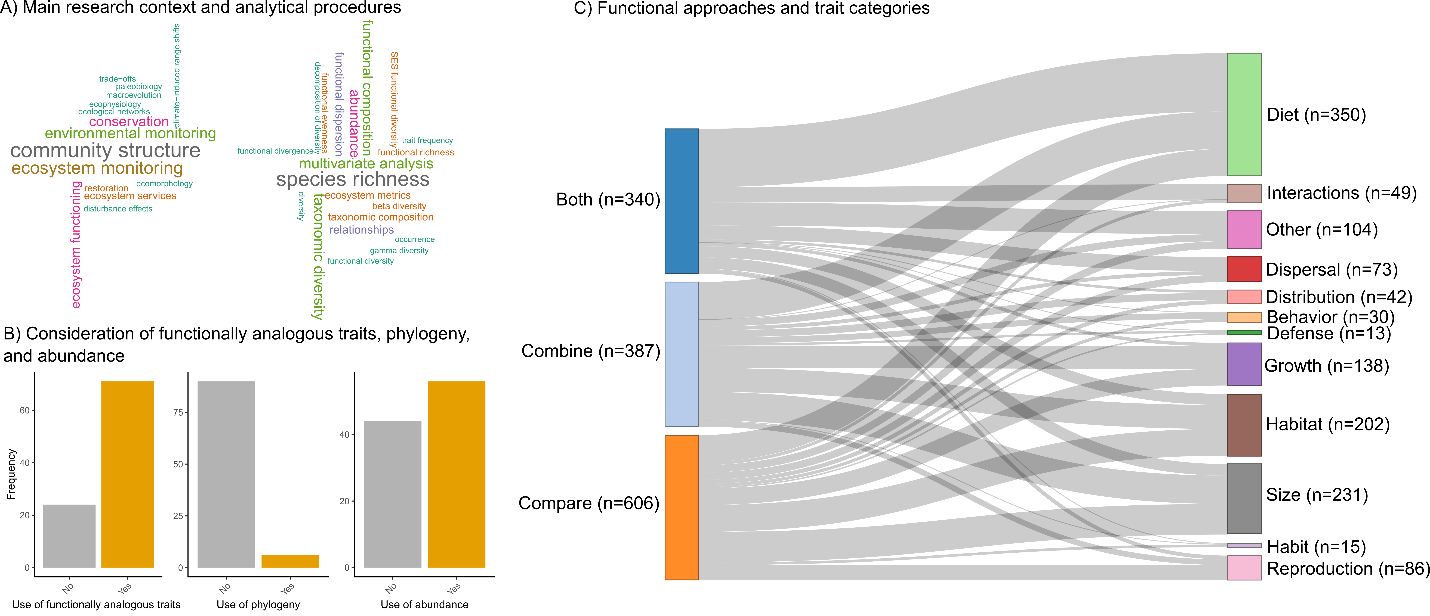
**Fig 1: PRISMA workflow showing the number of studied included and excluded along the systematic review process. The literature survey was conducted on Web of Science (WoS), on 08-06-2020.**



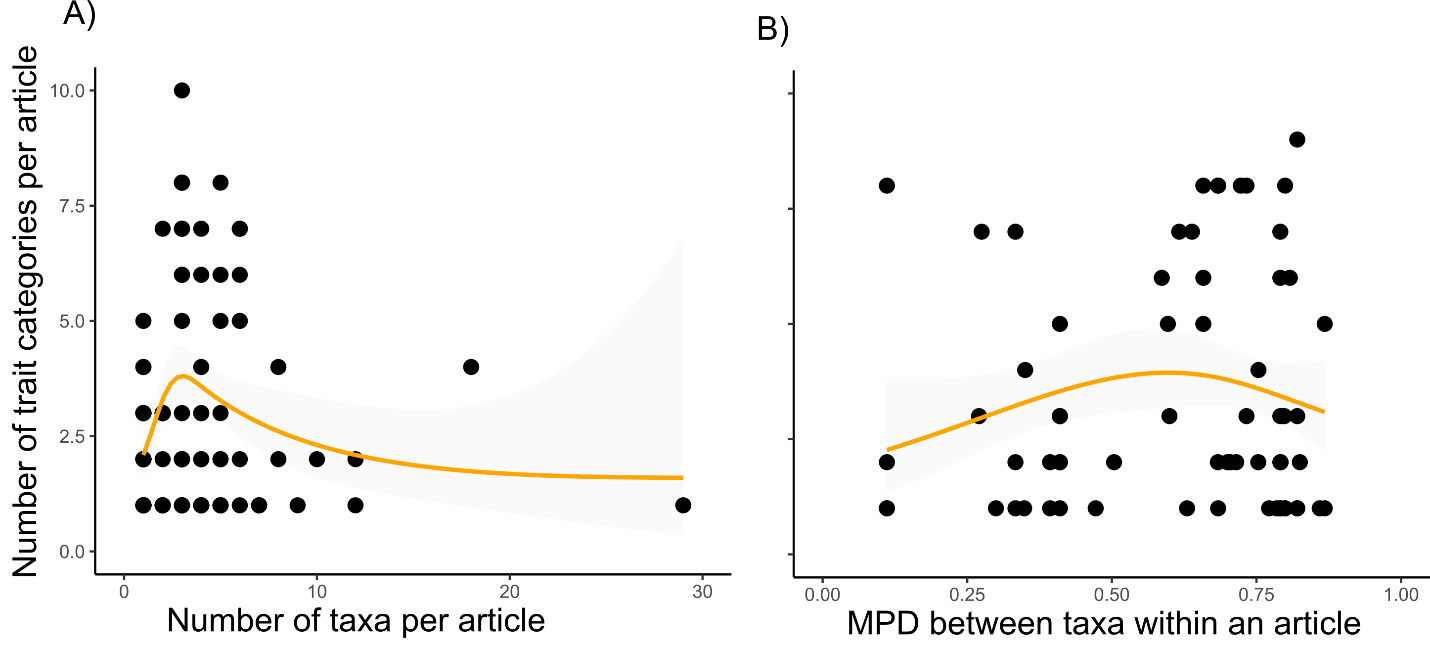
**Fig. 2: The identity of the taxonomic ranks (A), the number of taxonomic ranks (B), the frequency of each taxon (C) in cross-taxa research in functional ecology, D) Interaction graph showing the links between pairs of taxa in cross-taxa research in functional ecology. More than one taxon can be studied in the same article (see B)*,* so that the sum of values in A and C is larger than the total number of articles. In D, we present taxa at the level of class, the most often rank used in the surveyed articles, and link width refers to the number of studies per pair of taxa (values in the legend). The literature survey was conducted on Web of Science, on 08-06-2020.**



**Fig. 3: Topological tree of taxa comprising cross-taxa research in functional ecology. The phylogeny was based on the taxonomic ranks of organisms studied by the cross-taxa research (obtained using the function *classification* of *taxize* package of R). We used the ranks to build a matrix of taxa (rows) by taxonomic rank (columns); the cells in this matrix had either zeros or ones depending on whether a given rank belongs to a given taxon. We submitted this matrix to the Jaccard’s distance which calculates the distance between taxa based on the overall number of unique and shared taxonomic ranks. Finally, we applied the hierarchical clustering algorithm of complete linkage on the Jaccard distances to create the phylogeny. We colored the edges of the phylogeny by searching the edge (internal links in the phylogeny) between each pair of taxa, per article. We then summed the number of articles to obtain the total number of articles per edge. Jaccard distance was calculated using the function *vegdist* of the *vegan* package, and the complete linkage clustering was obtained using the function *hclust* of the *stats* package*,* both in R programming environment. Taxonomic nomenclature follows the NCBI (*National Center for Biotechnology Information*). The literature survey was conducted on Web of Science, on 08-06-2020.**



**Fig. 4: Main research contexts, analytical procedures and approaches used in cross-taxa research in functional ecology. In A we show word clouds (created by disassembling collected data) depicting the main research contexts and analytical procedures used in the research. The minimum frequency to appear in the cloud was set to two times. In B we show how often functionally analogous traits, phylogenies (as would be expected in phylogenetic contrasts or explicit use of evolution) and abundance are used in the research. In C we show a Sankey plot depicting the number of data entries (line width) that used different numbers of trait types (right labels) and functional approaches (combine cross-taxa data, compare cross-taxa data, both approaches in a same article). In C, we present the number of entries per functional approach and trait category (y axis of the left and right sides) because all bars in the left side have the same size after standardization (i.e., they sum 100%). An animated version of this Sankey plot can be found in the GitHub of the first author.**



**Fig. 5. The relationship between the number of trait categories and the number of taxa (A) and mean pairwise taxonomic distance (MPD) between taxa within an article (B). In B, only articles with at least two taxa were analyzed. The line and confidence intervals were estimated through a non-linear Poisson model with third order polynomial of the number of taxa and MPD. The shaded area depicts the 95% confidence interval around the non-linear trend.** **The literature survey was conducted on Clarivate Analytics Web of Science, on 08-06-2020.**

**Discussion**

Functional ecology research is much more concerned with ecological processes rather than patterns; they are generally addressed and inferred by interpreting variation in organismal traits (Bellwood et al. 2019, Tavares et al. 2019). The challenge of functional ecologists is to find emergent properties of whole communities and ecosystems by crossing the borders of taxonomy (Weiss & Ray 2019, Aubin et al. 2013).

We Movng from single to multiple taxa research ---

multiple to truly cross taxa

Such a concern was revealed by our data on the number of taxa, number of ranks, links between taxa, and coverage of taxa in the topological tree of life. Processes underlying community structure are the main focus of this research field, but also environmental factors that can trigger changes in ecological communities – as identified by the field of ecosystem/environmental monitoring. While comparing diversity across taxa and finding surrogate/indicator taxa to minimize sampling costs has been in the core of monitoring (e.g., Heino 2010, Landeiro et al. ?, Magg et al. 2019), we advocate that using an approach that combines taxa would lead us to an understanding of what is driving changes in whole ecosystems (further discussed below in “*A monitoring framework for functional and cross-taxa data*”).

***A monitoring framework for functional cross-taxa research***

Mouillot et al. (2013) recently formulated a framework that explicitly considers several facets of functional diversity while addressing the effect of disturbances on communities; as they did not explain how to apply to multiple taxa, we called it a “single taxon framework”. Mouillot and colleagues also explore biotic pressures on local communities. The effect of biological invasions, for instance, can extend farther from the effect on close relatives and spread across the whole ecosystem. By using a cross-taxa approach one can understand the response of multiple components to disturbances such as that one. Two factors are fundamental for extending this framework: 1) the taxonomic/ phylogenetic distance between taxa, and 2) consideration of abundance.

We found a trade-off between the number of taxa/ taxonomic distance and the number of trait categories possible to be addressed in research, meaning that the strength of inference will be influenced by how deep we dive into the tree of life. Overall, if we dive too superficially in the tree of life (i.e., within a taxon), we possibly will be addressing the fitness of one group of closely related organisms, rather than ecosystem properties (Fig. 6). In contrast, the deeper we dive in the tree of life the closer we are from understanding ecosystem properties (Fig. 6). However, be careful with the depth as we may find methodological and analytical difficulties that are difficult to resolve.

Abundance is a fundamental quantity in ecology and evolution (ref). Functional ecology also does a good use of this parameter by acknowledging that strength of ecological functions and interactions change with organismal abundance (ref. Benthos study, Mouillot et al. 2013). For instance ????. We think that abundance is still overlooked in cross-taxa research, and we will explain why we think that. The same ecological function can be performed in different regions by different organisms (what we call ‘ecological convergence’, e.g., Kelt et al. 1997, 2004), but it is hard to compare if the strength of a function differs across regions. One example of this is the function of granivory: while this function is performed by rodents in the deserts of North America, it is performed by ants in the deserts of South America. These organisms are related at the phylum level and therefore have no to few ecological attributes in common. To solve this we could measure, in both taxa, traits strongly correlated with the rate of consumption, transport and storage of grains (e.g.,), collate data as advised by Weiss & Ray’s framework, and then weight functional indices by taxon abundance in a set of geographically replicated sites. This way we could check whether traits related to granivory vary across regions.

Redundancy across taxa – ants of amazon seem to perform almost functions of that ecosystem.. how to discuss this factor?

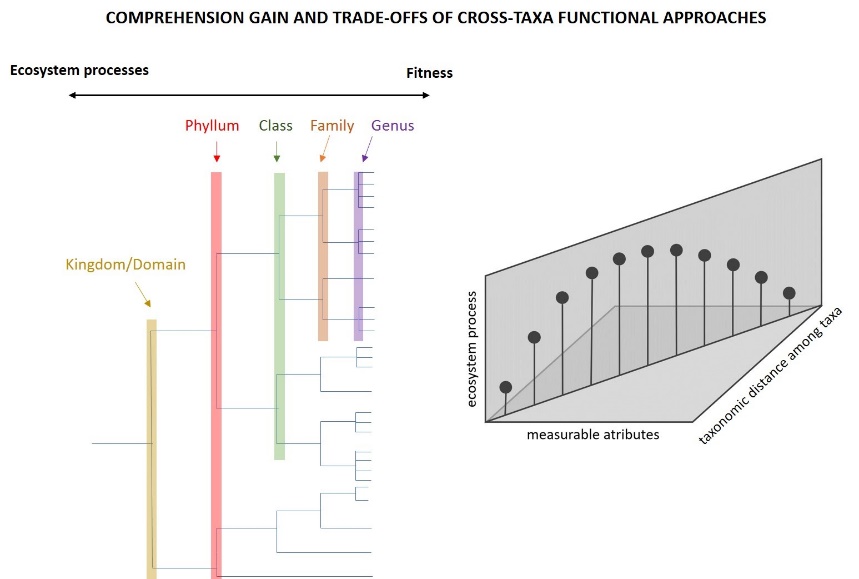


Fig. 6: Conceptual framework showing the processes being addressed as we dive deeper in the tree of life.

Functional ecology is moving from indexes based on dendrograms and distance matrices (eg Webb et al. 2002, Petchey & Gaston 2006) toward more sophisticated methods based on convex-hulls (Cornwell et al. 2006, Leprieur et al. 2008) and hypervolumes (Blonder et al. 2014, 2017) (reviewed by Mammola et al. 2021).

For example, one dominant research area is ecological indicator analysis (citation). If we have cross-taxa occurrence and trait data, we can elect the indicator taxa based on both local richness and functional diversity, and even elect the most functionally distinctive indicator taxa based on a combination of trait values.

We can also prioritize conservation considering multiple dimensions of biodiversity (Brum et al. 2017, Dobrovolski et al., Sobral et al.) and consider areas of importance considering species endemicity and functional distinctiveness.

And most importantly, we can evaluate which species or group is more critical to a given function and, finally, we can evaluate what combined taxa can do together relative to what can do in particular.

We advocate that using an approach combining cross-taxa data, with careful thinking about traits and the functions they represent over different taxa can aid ecology to advance the taxonomic frontiers of data, and go beyond by achieving the understanding of the functioning of whole communities and ecosystems (ecosystem components).

Concept:

Monitoram

Entendimento

Relacoes entre organismso

Pedro jordano – tamanho do fruto e tamanho do dispersor

**Good examples**

Froehlich et al. 2016 Aquaculture – nice to discuss

Mori et al. 2015 nice to discuss – dispersion traits plants and other

Bruno K.C. Filgueiras – congruence on insects - fragmentation

Discuss: https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.2924

Winsa et al. 2017

Combined vs. non-combined cross-taxa data

Ecosystem wide impacts of marine heat waves across marine and terrestrial ecosystems traits

– Ruthrof et al. 2018

Hypervolume --- facilitam a pespectiva cross-taxa em uma perspectiva de trait continuo

True cross taxa – cooke, pimiento, …

**Conclusion**

First, we advocate that the term ‘cross-taxa’ should be only used by research that crosses taxonomic boundaries, rather than researches that compare data of multi groups (which should use the term ‘multiple taxa’). Thus cross-taxa research can be defined as the ‘*research aiming to synthesize information of organism response to (fitness) and effect (function) on environment NECESSARILY involving distantly related and co-occurring taxa’*.

Esta revisão da literatura avança não somente na teoria sobre a pesquisa funcional *cross-taxa*, mas também mostra caminhos para migrar em busca da quantificação de propriedades emergentes e serviços ecossistêmicos através do uso de atributos quantitativos-chave e de fácil obtenção  (como tamanho, que pode ser obtido tanto para peixes como para bentos) em conjunção com atributos medidos de forma mais grosseira (e.g., fatores binários ou categóricos como a forma de obtenção de recursos alimentares e o hábitat ocupado).

On-line supporting information

Table S1: Questionnaire applied in the initial filtering phase. We describe the inclusion and exclusion criterion for each question.

|  |  |  |
| --- | --- | --- |
| Question / Inclusion | | Exclusion |
| Q1: Research in ecology? | |  |
|  | Research explicitly interested in testing the relationship between organisms and their environment. Most of studies were in the fields of community ecology and macroecology. \* | Research in knowledge fields such as physics, engineering, health sciences, economics. Research not directly interested in organismal-environment relationships, such as population genetics, phylogenetic systematics, taxonomy, historical biogeography, and paleontology. |
| Q2: Research in functional ecology? | |  |
|  | Research considering the relationship between traits and organismal performance and function (Arnold 1983, Violle et al. 2007), with potential implications for ecosystem services (Tavares et al. 2019). ‡ | Studies on trait evolution and ancestral character reconstruction, geometric morphometrics with focus on the mechanics of morphological structures lacking implication on community structure. Ecological networks research with focus on species rather than communities. |
| Q3: Cross-taxa research? | |  |
|  | Studies of organisms with a phylogenetic relationship at family and / or higher taxonomic ranks (e.g., clades: flowerless and flowering plants; classes: Mammalia, Aves; order: Carnivora; family: Cerambicidae). † | Studies of organisms with a phylogenetic relationship below family level |
| Q4: Literature review? | |  |
|  | Not a literature review. | Title and abstracts including the words “systematic review”, “review”, “bibliographic review” (as detailed in the PRISMA statement, Liberati et al. 2009). Data are of secondary importance in this group of study. |
| Q5: Meta-analysis? | |  |
|  | Not a meta-analysis | Title and abstracts including the word “meta-analysis” (as detailed in the PRISMA statement, Liberati et al. 2009). Data are of secondary importance in this group of study. |

\* We also included studies in paleoecology, phylogeography and biogeography, with explicit test of ecological processes underlying diversity, distribution, and evolution.

‡ Most studies were about community functional diversity at local and macroecological scales (e.g., ????). We also included studies using coarse functional entities/groups with clear functional reasoning behind their definition (generalist, specialist, terrestrial, arboreal) (e.g., Pardini et al. 2009), studies on trait-mediated ecosystem services (e.g., Werling et al. 2014, PNAS), and studies on geometric morphometrics comparing the morphological and ecological trait space filled by groups of species (e.g., herbivorous, carnivorous; e.g., ????).

† Q3 Some taxa are grouped by sampling convenience (e.g., macroinvertebrates, which includes crustaceans, insects, among others). In these cases we used the highest rank embracing all taxa.

The filtering process could be influenced by reader skills, as the filtering team includes one post-doc, one doctorate student, and two under-graduate students. To minimize such influence, we first ran a leveling phase (Fig. 1) where all readers answered the six questions based on information from the 20 most cited papers. After that, we met, checked the correspondence between answers, and solved doubts. Then, we started the filtering process by 1) ordering studies according to the first author’s surname —to avoid that only one reader reads the most cited and complex studies—, 2) dividing the 1,006 studies among the four readers, 3) answering the questions (Table 1). Doubts that arose during the filtering process were registered in an online document open to the team of readers, and were resolved by AL Luza (the most experienced reader in the filtering team). Answer options were ‘NO’, ‘YES’, ‘MAYBE’, and ‘NA’. The answer ‘MAYBE’ was used when information from title and abstract were not enough to reach ‘YES’ or ‘NO’. The answer ‘NA’ was used when none answer was applicable.

Objective 1

After completely solving the ‘MAYBE’ answers -- through a quick read on the main text--, we run a final filtering cased on positive responses of Q1 to Q5 to achieve our first objective. We consider systematic reviews and meta-analysis as they can provide values and analysis, as well valuable information about available approaches to cross-taxa studies.

Articles with either positive or negative responses to functional ecology were called, respectively, ‘functional ecology research’ and ‘general ecological research’ area. Such differentiation is important because general ecological research (mostly community ecology and macroecology) assumes all taxa as ecologically different, while functional ecology explicitly considers degrees of ecological differences between taxa (Fig. 1).

Table S2: Complete set of descriptors collected in the surveyed literature.

|  |  |
| --- | --- |
| **Descriptor** | **Description** |
| PaperNumber | The number of the paper, among the 1,006 articles found in WOS survey |
| YourName | Complete name of the reader (e.g, André Luís Luza) |
| StudyAuthors | List of authors of the study being read |
| StudyTitle | Title of the study being read |
| Is it really a cross-taxa research in functional ecology? | We defined 'family' as the minimum taxonomic level to be cross-taxa (as niche conservatism used to be low above family level) |
| Are functional aspects based on traits or habitat affinities | Whether functional perspective is based on traits or rough definitions of trait (often habitat) |
| Does it declare to be ‘cross-taxa’ research? (In title, abstract, along the text) | Many studies are cross-taxa, but have no stated so. We would like to know how many articles omit such information. |
| Position of first citation of 'cross-taxa' terms ("cross-taxa", "multi-taxa", "multiple-taxa", "taxon-free") | Does the study immediately (i.e., in the title) claim to be cross-taxa? |
| Type of study | We would like to know what is the major type of study using data across taxa |
| Research context | We would like to know what is the broad ecological context in which cross-taxa data are applied |
| Detailed research context | We would like to know a more detailed context in which these data are applied |
| Key message | The take-home message (if possible, summarize it in one sentence) |
| Time-series? | Are cross-taxa data collected over time? |
| Which organisms? | Organisms under study (major groups) |
| Taxonomic rank | Taxonomic level embracing the taxa being studied. Fill with The most superior level ("Division, phylum") not the inferir ("spp") |
| Which traits? | Traits used to characterize the functional structure of ecological assemblages |
| Is it an easily obtained 'soft' trait? | State if each trait is a soft (easily obtained) or 'hard trait' (hard to get) |
| Were quantitative trait values used (yes=quantitative; no=ordinal, category, binary)? | We would like to know whether the traits used are quantitative or not |
| Were abundance data used? | If counting of individuals, estimated abundance, were used in the functional analyses |
| How do authors use cross-taxa data? | We would like to know whether authors used to compare data across taxa (e.g., they run one analysis per taxon), combine data across taxa (e.g., the gather all data to estimate a global richness, functional diversity), or both alternatives |
| Do they use the same traits across taxa (are traits functionally analogous across taxa)? | We would like to know whether authors used traits that describe similar ecological functions and processes across taxa |
| Which index and analytical procedures were used? | What kind of index and/or analyzes did the authors use? |
| Do they use phylogeny? | We would like to know how many studies explore an evolutionary perspective |
| Which ecological system? | Major ecological system under study |
| Which habitat type? | Major habitat type under study |
| Which realm? | Terrestrial and marine realm under study |

Table S3: List of terms included in the 12 diet categories used in this study. All these terms were mentioned by the authors that screened the articles (see the section of authors’ contribution).

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Growth | Diet | Habitat | Size | Reproduction | Dispersal | Defense | Distribution | Interactions | Behavior | Habit | Other |
| growth\* | diet | habitat | size | reproduction | migrat\* | thorns | geographic\* | \*host\* | activit\* | semi-fossorial | none |
| wood density | feeding | Habitat | weight | nesting\* | dispersal | wood density | distribution\* | competitive\* | overwinter strategy | sociability | functional types |
| storage organs | foraging | shade\* | height | nest\* | dispersal | \*compounds | red listed species | parasitoid\* | home range | sociality | functional group (small protists, large protists, mixotrophs) |
| trunk xylem density\* | food | specializat\* | heigh | life\* | balloning | color | conservation status | flower visitor | escape\* | hunting\* | market value |
| leaf\* | guild | strata | mass | reproductive\* | ballooning | color | origin | flowering | diurnal | prey capture\* | current preference |
| foliar | trophic | light\* | basal area | asexual\* | walk\* | defence | invasibility | pollinat\* | mean laying date | eye width | ecological strategy |
| temperature\* | herbivores | stratum | volum\* | prothallus\* | locomotion | palatability | endemism | social organization | active period | visual acuity | taxa classes |
| \*thermal\* | hypsodonty | soil position | length | spore\* | \*chorous | trichome | rarity | pollen | adult active | | plant type |
| metabolism\* | predators | disturbance | body ratio | fecundity | conidia | spine | native | number of queens | |  | abundance trends |
| climatic\* | omnivores | forest | shape | voltinism | wing\* | resistance | historic |  |  |  | trend in abundance |
| niche\* | detritiv\* | aridity score | thallus form | \*basidiomes | mobility | mimicry |  |  |  |  | surface |
| tolerance\* | tooth\* | brachial\* | form | ascomata | dissemination\* | |  |  |  |  | ant |
| moisture\* | lower jaw morphometrics | sensillus\* | stem diameter | breeding\* | movement | |  |  |  |  | non-ant |
| humidity requirement | bill lenght | humidity preference | intertegular | fruit\* | aerial |  |  |  |  |  | functional type |
| morphology\* | setae\* | substrate |  | galertic\* | dispersed | |  |  |  |  | surface structure |
| torpor | photobiont\* | sucessional\* | | propagation predominant vegetative | dispersion syndrome | | |  |  |  | category |
| lateral spread | oxygenic photosynthetic | saproxylic\* | | seed bank | |  |  |  |  |  |  |
| nitrogen-fixing | jaw | water dependence | | spawning preference | | |  |  |  |  |  |
| laminar surface area | microwear | zonation |  | legumino |  |  |  |  |  |  |  |
| perennial | buccolingual | altitude |  |  |  |  |  |  |  |  |  |
| number of claws | curvature | hydrologic | |  |  |  |  |  |  |  |  |
| \*root\* | enamel | ph preference | |  |  |  |  |  |  |  |  |
| hypha | mesiodistal | plant parts eaten | |  |  |  |  |  |  |  |  |
| foliag\* | occlusal | plant parts and | |  |  |  |  |  |  |  |  |
| foliage persistance | lecty | hypogeic |  |  |  |  |  |  |  |  |  |
| cellulose percentage |  | wetland use | |  |  |  |  |  |  |  |  |
| vertical stratification |  |  |  |  |  |  |  |  |  |  |  |
| structural complexity |  |  |  |  |  |  |  |  |  |  |  |
| laminar total chlorophyll |  |  |  |  |  |  |  |  |  |  |  |
| fiber percentage |  |  |  |  |  |  |  |  |  |  |  |
| carbon percentage |  |  |  |  |  |  |  |  |  |  |  |
| crown width |  |  |  |  |  |  |  |  |  |  |  |
| flexibility |  |  |  |  |  |  |  |  |  |  |  |
| minimum dissolved oxygen |  |  |  |  |  |  |  |  |  |  |  |
| head fraction |  |  |  |  |  |  |  |  |  |  |  |
| plant functional type |  |  |  |  |  |  |  |  |  |  |  |
| colony type |  |  |  |  |  |  |  |  |  |  |  |
| ability to overgrow other colonies |  |  |  |  |  |  |  |  |  |  |  |
| nutrient uptake |  |  |  |  |  |  |  |  |  |  |  |
| ability to capture p |  |  |  |  |  |  |  |  |  |  |  |
| ability to degradate c |  |  |  |  |  |  |  |  |  |  |  |
| tiller |  |  |  |  |  |  |  |  |  |  |  |
| longevity |  |  |  |  |  |  |  |  |  |  |  |
| duration |  |  |  |  |  |  |  |  |  |  |  |
| larval development |  |  |  |  |  |  |  |  |  |  |  |
| stages |  |  |  |  |  |  |  |  |  |  |  |
| respiration |  |  |  |  |  |  |  |  |  |  |  |

Table S4: Frequency in trait categories across data entries. Values per trait category are number of entries standardized by the number of entries per functional approach.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Functional approach | Trait category | | | | | | | | | | | | |
|  | Behavior | Defense | Diet | Dispersal | Distribution | Growth | Habit | Habitat | Interactions | Other | Reproduction | Size |
| Both | 0.29 | 0.29 | 40.29 | 9.71 | 1.76 | 2.35 | 0.59 | 7.65 | 10.59 | 15.88 | 2.65 | 7.94 |
| Combine | 4.13 | 1.55 | 25.84 | 2.58 | 4.91 | 15.50 | 0.52 | 16.54 | 0.00 | 4.91 | 3.88 | 19.64 |
| Compare | 2.15 | 0.99 | 18.65 | 4.95 | 2.81 | 11.55 | 1.82 | 18.48 | 2.15 | 5.12 | 10.23 | 21.12 |

Table S5: Model selection analysis showing the ranking of models tested to find a trade-off between number of taxa (or MPD) and number of trait categories per study.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Predictor variable/ Model | | | Intercept (log scale) | Coefficient | df | logLik | AICc | deltaAIC | weight |
| Number of taxa | |  |  |  |  |  |  |  |  |
|  | Third order polynomial | | 1.114 | + | 4 | -209.78 | 428 | 0 | 0.641 |
|  | Second order polynomial | | 1.12 | + | 3 | -211.838 | 429.9 | 1.94 | 0.244 |
|  | First order polynomial | | 1.165 | -0.009 | 2 | -213.653 | 431.4 | 3.43 | 0.115 |
| Mean Parwise Distance (MPD) | | |  |  |  |  |  |  |  |
|  | Second order polynomial | | 1.254 | + | 3 | -156.471 | 319.3 | 0 | 0.617 |
|  | Third order polynomial | | 1.254 | + | 4 | -156.466 | 321.6 | 2.26 | 0.199 |
|  | First order polynomial | | 1.03 | 0.391 | 2 | -158.782 | 321.8 | 2.43 | 0.183 |

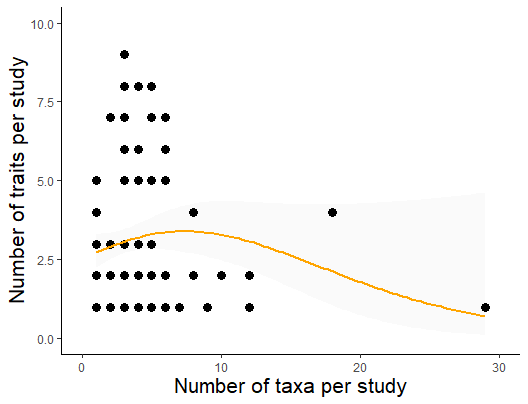


Fig. S1. **The relationship between the number of trait categories and the number of taxa. The line and confidence intervals were estimated through a non-linear Poisson model with a second order polynomial of the number of taxa. The shaded area depicts the 95% confidence interval around the non-linear trend.** **The literature survey was conducted on Clarivate Analytics Web of Science, on 08-06-2020.**