Date: December 2023

Author: Pereira, Catia L.

OXFORD

BioScience, 2023, 73, 862–878

https://doi.org/10.1093/biosci/biad089

Advance access publication date: 1 November 2023

Overview Article

Future-proofing environmental DNA and trait-based predictions of food webs

Cátia Lúcio Pereira, Zeynep Ersoy, M. Thomas P. Gilbert, Dominique Gravel, Miguel B. Araújo and Miguel G. Matias

Cátia Lúcio Pereira (catiaalpereira@gmail.com) is a postdoctoral researcher in aquatic ecology and affiliated with Rui Nabeiro Biodiversity Chair, MED – Mediterranean Institute for Agriculture, Environment and Development & CHANGE Global Change and Sustainability Institute, Institute for Advanced Studies and Research, University of Évora, in Évora, Portugal, Globe Institute, University of Copenhagen, in Copenhagen, Denmark and Departamento de Biogeografía y Cambio Global, Museo Nacional de Ciencias Naturales (MNCN-CSIC), in Madrid, Spain. Zeynep Ersoy (zzeynepersoy@gmail.com) is a postdoctoral researcher in aquatic ecology and affiliated with Rui Nabeiro Biodiversity Chair, MED – Mediterranean Institute for Agriculture, Environment and Development & CHANGE Global Change and Sustainability Institute, Institute for Advanced Studies and Research, University of Évora, in Évora, Portugal, FEHM-Lab (Freshwater Ecology, Hydrology and Management), and Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona (UB), in Barcelona, Spain. M. Thomas P. Gilbert (tgilbert@sund.ku.dk) is a Professor at the Globe Institute, University of Copenhagen, in Copenhagen, Denmark and is also affiliated with University Museum, NTNU, Trondheim, Norway. Dominique Gravel (dominique.gravel@usherbrooke.ca) is a Professor at the University of Sherbrooke, Québec, in Canada. Miguel B. Araújo (maraujo@mncn.csic.es) is a Research Professor at the Departamento de Biogeografía y Cambio Global, Museo Nacional de Ciencias Naturales (MNCN-CSIC), in Madrid, Spain, and also a Professor at the Rui Nabeiro Biodiversity Chair, MED – Mediterranean Institute for Agriculture, Environment and Development & CHANGE Global Change and Sustainability Institute, Institute for Advanced Studies and Research, University of Évora, in Évora, Portugal. Miguel G. Matias (miguel.matias@mncn.csic.es) is a Ramón y Cajal Research Fellow at Departamento de Biogeografía y Cambio Global, Museo Nacional de Ciencias Naturales (MNCN-CSIC), in Madrid, Spain and a Visiting Researcher at the Rui Nabeiro Biodiversity Chair, MED – Mediterranean Institute for Agriculture, Environment and Development & CHANGE Global Change and Sustainability Institute, Institute for Advanced Studies and Research, University of Évora, in Évora, Portugal. Cátia Lúcio Pereira and Zeynep Ersoy contributed equally to this work.

Abstract

Food webs represent trophic interactions within ecosystems. Matching traits of consumers and resources helps infer trophic interactions and food-web properties. Environmental (e)DNA, commonly used for detecting species occurrences, is rarely used in trait-matching studies because abundance estimates and descriptions of relevant traits are generally missing. We synthesized recent literature on inferences of trophic interactions with eDNA and trait matching to identify challenges and opportunities for coupled eDNA–trait recording schemes. Our case study shows how coupling eDNA and trait data collection improves the ability to characterize greater numbers of food webs across multiple scales ranging from spatiotemporal to trait variation. Future-proofing eDNA data sets requires the collection of new traits or the compilation of existing trait data at spatiotemporal scales that are relevant to detect current and future changes in food webs and ecosystems.

Keywords: environmental DNA, trait matching, food webs, networks, functional traits

Food webs represent energy flows between species or groups of species (Barnes et al. 2018). They provide a framework for examining the roles of species within ecosystems and determine the mechanisms through which biodiversity influences ecosystem functioning (Evans et al. 2016, Barnes et al. 2018). It is often difficult to characterize food-web structures empirically (Evans et al. 2016), because large-scale sampling of biodiversity remains unachievable for most ecosystems and because the existing samples of trophic interactions are uneven and incomplete (Mestre et al. 2022a, Ficetola and Taberlet 2023). Moreover, once sampling is achieved, node attributes (e.g., species, guilds, body sizes) need to be linked according to their trophic relationships, a process plagued by a variety of uncertainties. Despite the formidable size of the problem (Morales-Castilla et al. 2015), recent advances in environmental DNA (eDNA) have opened new possibilities for addressing this challenge. eDNA refers to the DNA that organisms shed into their surroundings, which can be extracted from environmental samples (figure 1a, box 1; Taberlet et al. 2012). By employing metabarcoding techniques and trait-matching approaches (comparing traits of potentially interacting species to predict the likelihood of their interaction; figure 1b, box 1; Laigle et al. 2018), combined with parametric models such as generalized linear models (Gravel et al. 2013, Pomeranz et al. 2019) or machine learning algorithms such as random forest or k-nearest-neighbor (Laigle et al. 2018), the field of ecosystem ecology can leverage existing and emerging data to enhance our understanding of species interactions and unravel intricate food webs.

There are, however, biases associated with molecular-based approaches when inferring biotic interactions with trait matching. For example, obtaining estimates of species' abundance, biomass, and traits using eDNA is far from straightforward (Derocles et al. 2018, Clare et al. 2019, Siegenthaler et al. 2019). The mismatch between eDNA and available trait data sets due to the independent collection of species diversity, traits, and interactions (Compson et al. 2018, Pellissier et al. 2018) and the associated limitations cannot be ignored (e.g., incomplete databases, inconsistent taxonomic resolution; McGee et al. 2019, Pomeranz et al. 2019). Being aware of these challenges from the beginning of experimental design will help mitigate several associated issues through well-designed recording schemes built around the research question at hand (Cuff et al. 2022). eDNA represents a step forward in the increasing field of DNA monitoring (D'Alessandro and Mariani 2022).

Received: January 26, 2023. **Revised:** August 28, 2023. **Accepted:** September 12, 2023

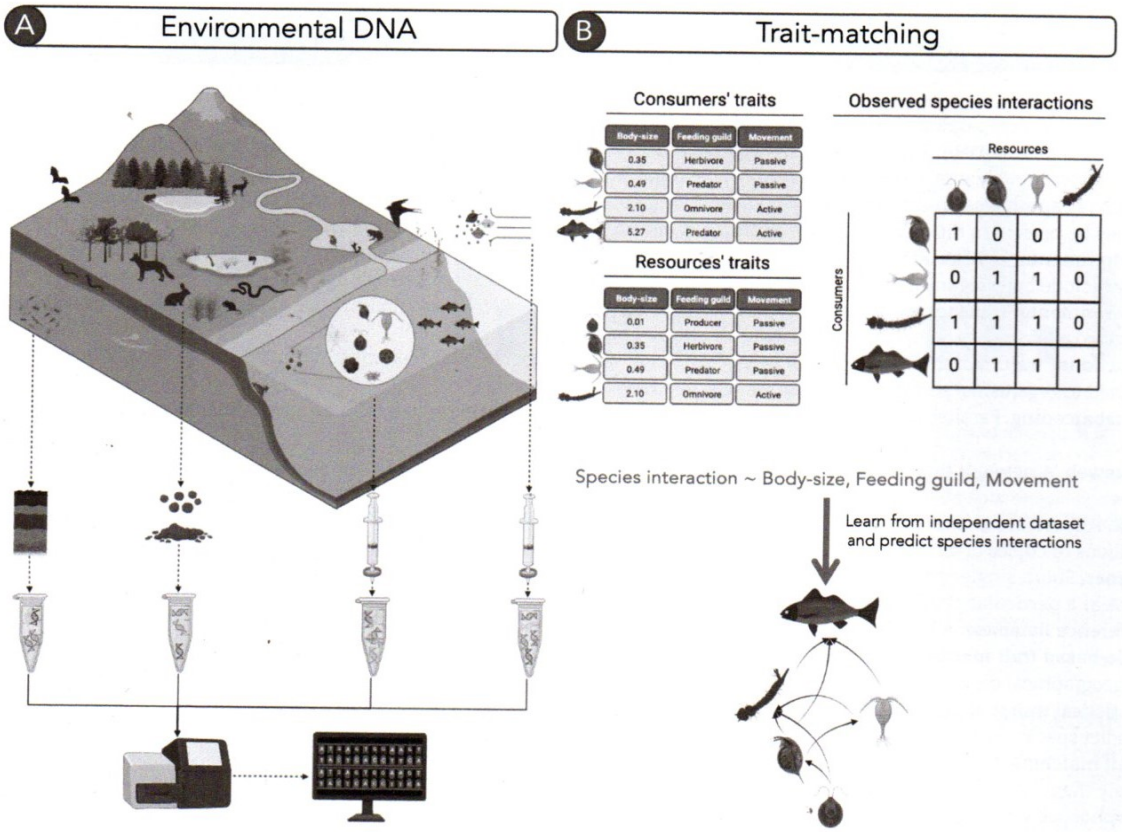


Figure 1. Conceptual illustration of environmental DNA and trait-matching approaches. (a) Environmental DNA workflow. eDNA is isolated from an environmental sample (sediments, water, or air) and subsequently sequenced (shown as Illumina MiSeq technology). Finally, high-throughput sequencing data are computationally processed (bioinformatics), where data are quality filtered, summarized, and compared with reference databases for taxonomic assignment. (b) Trait matching. The statistical algorithm aims to estimate the probability of a consumer–resource interaction, on the basis of their trait values, and infer interaction networks. Illustration created with BioRender.com.

offering opportunities to untie diet studies from the constraints of collecting animals from wild populations (Clare 2014).

Despite several efforts for establishing a framework to reconstruct food webs (e.g., Bohan et al. 2017, Compson et al. 2018, Derocles et al. 2018, Dubart et al. 2021), to our knowledge, there is no comprehensive framework to infer food webs combining eDNA data and highly resolved trait information. In the present article, we review recent literature using DNA metabarcoding or trait-based approaches to infer food webs; synthesize challenges and limitations of each approach; propose a framework to couple eDNA and trait data to reconstruct food webs, showcasing how it improves our capacity to characterize food webs across scales using empirical data from a multiregion experimental facility of freshwater pond mesocosms; highlight a set of functional traits to improve food-web predictions; and provide recommendations toward integration of eDNA–trait recording schemes for improved detection of changes in food-web structures, while ensuring these data sets remain relevant in the future.

Opportunities of using eDNA and trait matching to infer trophic interactions

Molecular approaches have experienced a massive surge of popularity in ecology. The capture and analysis of eDNA allows the generation of biodiversity data with noninvasive, cost-effective, and whole-ecosystem surveys (Taberlet et al. 2012, Bohmann et al. 2014, Seymour et al. 2020, Duarte et al. 2021), thereby enhancing sampling repeatability and species detection and identification (Taberlet et al. 2012, Bohmann et al. 2014, Cantera et al. 2019). This, in turn, allows biodiversity monitoring efforts to be scaled up, particularly in complex and large systems (Blackman et al. 2022). eDNA methods offer a unique perspective on biodiversity by simultaneously studying organisms across multiple trophic levels and domains of life (Djurhuus et al. 2020), thereby providing insights into complex biotic interactions and ecosystem change across space and time (Vacher et al. 2016, Compson et al. 2019, D'Alessandro and Mariani 2021). The expansion of eDNA has opened new research opportunities for studies on diverse ecological interactions, including pollination (e.g., for a review, see Banerjee et al. 2022), mutualism (e.g., Rasmussen et al. 2021), and trophic interactions (e.g., for reviews on species diets, see Clare 2014 and Alberdi et al. 2019; for co-occurrence food webs, see Seymour et al. 2020, Blackman et al. 2022). This indicated the broad applicability of eDNA as a multidisciplinary approach (Deiner et al. 2021, Veilleux et al. 2021) that can address complex ecological questions related to interactions between different taxa.

Morphology-based approaches that provide relevant trait data (Duarte et al. 2021), abundance (Olivier et al. 2019, Pomeranz et al. 2019), and complementary information about community

Box 1. Glossary.

Bioinformatics. Computational processing of sequence data, in which high-throughput sequencing data are quality filtered, summarized, and compared with reference databases for taxonomic assignment.
Environmental (e)DNA. DNA, potentially degraded, isolated from an environmental sample (e.g., water, sediment, air, soil). May include both organismal DNA, derived from whole organisms in the sample (e.g., bacteria, microalgae), and extraorganismal DNA, which is captured separately from the organism it originated from (e.g., cells, organelles, shed skin, scales, feces, saliva, gametes). It may be present in different states, for example, membrane-bound, adsorbed to particles, and dissolved.
Environmental (e)RNA. RNA isolated from an environmental sample, captured separately from the organism it originated from. May include both organismal RNA (e.g., prokaryotes), and extraorganismal RNA as cellular, vesicular, or free form.
Environmental transcriptomics. Gene expression profiling based on eRNA. It is a noninvasive approach that does not require prior knowledge of species composition, spanning all trophic levels.
Functional trait. Morphological, physiological or phenological property of an organism, measurable at the individual level and related to organisms' performance or an ecological process.
Metabarcoding. Parallel sequencing of DNA barcodes using the total DNA extracted from an environmental sample (e.g., soil, water, air).
Metaweb. A network that includes all species occurrences and potential trophic interactions at any given time and site within an area.
PCR. Polymerase chain reaction is a method that uses thermal cycling in the presence of a polymerase enzyme to rapidly create millions of copies of a DNA fragment.
Primer. Short, single-stranded nucleic acid molecule consisting of a sequence of DNA bases that are designed to match the target DNA at a particular point in the genome.
Reference database. A library of DNA sequences derived from specimens of known identity.
Rule-based trait matching. Application of trait matching using special rules/hypotheses and information from traits, phylogenies or geographical distributions to define the interaction probabilities and identify the forbidden links.
Statistical trait matching. Use of machine learning algorithms (e.g., random forest, boosted regression trees, *k*-nearest-neighbor) to predict species interactions on the basis of a set of species' traits.
Trait matching. Comparison of traits (e.g., body size) of potentially interaction pairs of species (or functional groups) to predict how likely they are to interact.
Trophospecies. Group of taxa that share similar resources and consumers.

composition (Keck et al. 2022) can be useful when inferring species interactions across multiple scales (González-Varo and Traveset 2016). Functional traits (e.g., body size, feeding guild, movement type) are often used as predictors of trophic interactions and ecosystem functioning (McGill et al. 2006), because they can influence both community dynamics and ecosystem processes and, in turn, can be affected by community structure (Maureaud et al. 2019). Intra- and interspecific trait variation can affect species coexistence (Turcotte and Levine 2016) and ecosystems' abilities to cope with environmental changes (González-Suárez and Revilla 2013) and control species' interactions (Zhao 2014). The underlying hypothesis is that species interact when their functional properties (or traits) are compatible, making interactions possible (Eklöf et al. 2013). Trait-matching models have been increasingly used to infer ecological networks across different systems, including terrestrial (Laigle et al. 2018, Brousseau et al. 2018b, Mendoza and Araújo 2019, 2022), freshwater (Gray et al. 2015, Compson et al. 2018, Pomeranz et al. 2019), and marine ecosystems (Albouy et al. 2019, Pecuchet et al. 2020). These models use compatibility (or match) to predict the probability of an interaction occurring between two species. In broad terms, there are two families of trait-matching models (Laigle et al. 2018, Pichler et al. 2020): phenomenological or recommender models (e.g., *k*-nearest-neighbor; Desjardins-Proulx et al. 2017) and hypothesis based (e.g., matching optimal predator and prey body sizes; Gravel et al. 2013, Pomeranz et al. 2019). Thanks to recent advances in machine learning techniques and increasing computing power, researchers can leverage these methods to make more accurate predictions of ecological networks (Pichler et al. 2020, Strydom et al. 2021, Pichler and Hartig 2023), because they are more flexible, perform better, and can achieve high accuracy in predictive and classification tasks within a short time.

Overview of inferring food webs with eDNA and trait matching

Despite the added value of eDNA and trait-based approaches, there are some limitations for reconstructing food webs that need being recognized and addressed.

Literature review

To understand the limitations associated with inferences of trophic interactions using eDNA, DNA metabarcoding, or trait matching, we conducted a systematic search of studies published between 2015 and 2022 in Scopus. We constructed our query by including variations of our terms of interest—namely, DNA metabarcoding, species interactions, and trait matching (further details are available in supplemental file S1). Given the widespread use of DNA metabarcoding in the study of dietary composition (Bohmann et al. 2014, Clare 2014, Alberdi et al. 2019, Clare et al. 2019), we extended our review to include these studies. We reasoned that this inclusion was justified because these studies face similar challenges related to sample and data processing. The literature search resulted in 540 studies (DNA metabarcoding, 328; trait matching, 212; supplemental figure S1), and after searches of reference lists, missing references were identified and

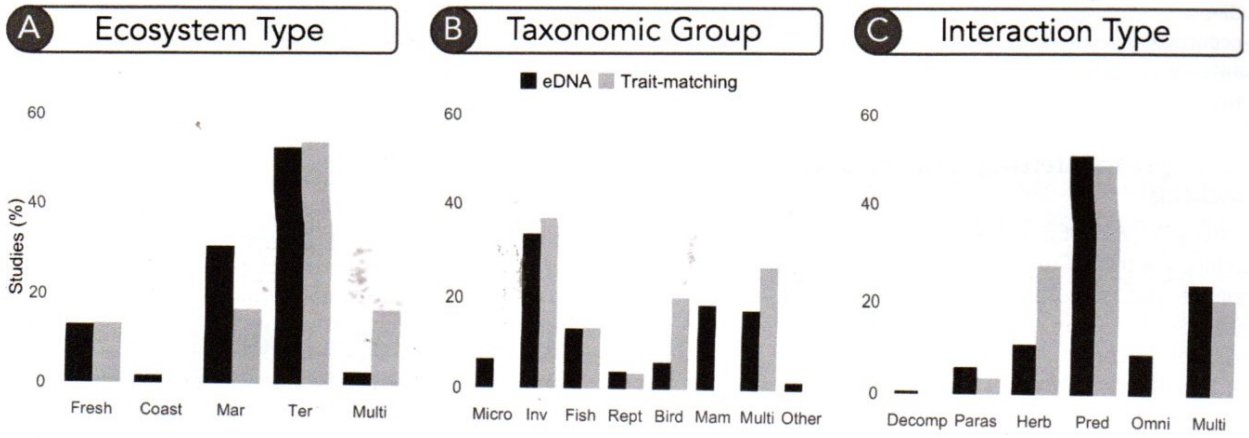


Figure 2. Literature review. The percentage of studies using DNA metabarcoding or trait matching to infer trophic interactions as a function of (a) ecosystem type, (b) taxonomic group, and (c) interaction type. The multiple category refers to studies that were focused on more than one ecosystem type, taxonomic group, or interaction type. Studies that did not specify a particular ecosystem type (eDNA, 17; trait matching, 6), taxonomic group (eDNA, 10; trait matching, 6), or interaction type (eDNA, 9; trait matching, 7) were excluded (e.g., overall reviews, comments, or opinions on the topic). Abbreviations: coast, coastal; decomp, decomposition; fresh, freshwater; herb, herbivory; inv, invertebrates; mam, mammals; mar, marine; micro, microorganisms; multi, multiple; omni, omnivory; paras, parasitism; pred, predation rept, reptiles; ter, terrestrial.

added manually (DNA metabarcoding, 21; trait matching, 6). A total of 567 studies were screened, and 244 were within the scope of our review (DNA metabarcoding, 208; trait matching, 36).

A comprehensive review of studies using DNA metabarcoding or trait matching to infer trophic interactions revealed a predominant focus on terrestrial ecosystems (52% and 53%, respectively) as the primary ecosystem type (figure 2a). Marine ecosystems followed terrestrial ecosystems in studies using DNA metabarcoding (30%) and trait matching (17%). Among the taxonomic groups, invertebrates were the most studied group by DNA metabarcoding (33%) and trait matching (37%; figure 2b). Although mammals were the second most studied group in DNA metabarcoding (19%), none of the trait-matching studies was focused exclusively on mammals, whereas studies on birds were more common (20%). In both approaches, a significant proportion of studies targeted multiple taxonomic groups (DNA metabarcoding, 17%; trait matching, 27%). Among the different types of interactions examined (figure 2c), predator–prey interactions were the most studied by both approaches (DNA metabarcoding, 51%; trait matching, 48%). Herbivory was the next most prominent interaction type in trait matching (28%), whereas DNA metabarcoding studies addressed multiple interaction types (21%). The literature review also showed that DNA metabarcoding has been applied across a wider range of ecosystems, taxonomic groups, and interaction types than trait matching has. This indicates the growing interest in reconstructing ecological networks using taxonomic data generated by eDNA metabarcoding (e.g., Bohan et al. 2017, Derocles et al. 2018, Dubart et al. 2021). However, only a few studies integrated DNA metabarcoding with trait-based methods for food-web reconstruction (Hardy et al. 2017, Compson et al. 2018, 2019, Djurhuus et al. 2020, Foster et al. 2020, D'Alessandro and Mariani 2021, Blackman et al. 2022, Pitteloud et al. 2022). These studies highlight that, despite the limitations of DNA metabarcoding and trait-matching approaches, their combined use through the integration of biodiversity recording and monitoring schemes using eDNA coupled with trait-based approaches (hereafter, *eDNA–trait recording schemes*) can provide a more comprehensive understanding of complex food webs. Pairing such data can improve overall network completeness and facilitate the inclusion of taxa otherwise methodologically overlooked (Cuff et al. 2022), revealing the potential missing links.

Challenges in inferring food webs with eDNA

The main challenges of using eDNA to reconstruct food webs are related to DNA collection and laboratory processing, taxa detection and taxonomic assignments, and data resolution. Most of these limitations are methodological and include a lack of standardized laboratory and bioinformatics protocols (McGee et al. 2019, Duarte et al. 2021) and biases in DNA extractions and primers (Albaina et al. 2016, Derocles et al. 2018). Moreover, taxonomic assignment accuracy relies on the coverage of reference databases, which varies across taxa (Albaina et al. 2016, McGee et al. 2019) and causes different levels of taxonomic resolution (Albaina et al. 2016, Clare et al. 2019). Furthermore, eDNA does not allow to distinguish whether an organism is alive or dead (Albaina et al. 2016), and dead organisms often shed large amounts of DNA into the environment (Yates et al. 2021). Therefore, the origin, degradation rates, and persistence of DNA in the environment also pose challenges for accurate identifications (Barnes et al. 2014, Shogren et al. 2018, Siegenthaler et al. 2019). All of this may lead to high rates of false positives (i.e., the misdetection of species that are not present) and false negatives (i.e., undetected species that are present; Derocles et al. 2018), creating additional biases across ecosystems. Acknowledging and taking into consideration potential detection biases is crucial to minimize introducing further biases in food-web inferences.

For species to interact, they need to co-occur in both space and time. However, eDNA offers a snapshot of species occurrences in a specific location in a short time frame, showing a very heterogeneous time coverage (Harrison et al. 2019); therefore, links that do not exist might be predicted. Furthermore, despite several efforts (for reviews, see Ruppert et al. 2019, Siegenthaler et al. 2019), reliable quantitative estimates (abundance and biomass) of the surveyed species with eDNA are still not straightforward (Derocles et al. 2018, Clare et al. 2019, Siegenthaler et al. 2019). In addition, although data on the presence or absence of species provide a starting point to examine the set of potential interactions in any given location (Araújo et al. 2011, Cazelles et al. 2016),

studies are still enquiring as to what information content these co-occurrences reveal and whether they link to underlying community interactions, including trophic interactions (e.g., Galiana et al. 2023).

Challenges in inferring food webs with trait matching

According to the surveyed literature, the major limitations of trait matching are related to data availability, data quality and resolution (i.e., node resolution, type and resolution of the traits; for reviews, see Schneider et al. 2019, Mestre et al. 2022a), and the method of inferring the interactions (supplemental file S1). Trait data are often insufficient (i.e., we are unable to measure the traits determining an interaction) or of low quality (e.g., low spatiotemporal resolution). Most traits are not directly measured but are available at coarse resolutions (both spatiotemporal and taxonomic), grouped in bins (e.g., size classes), or derived from expert knowledge (e.g., guilds), and several taxonomic groups are often resolved at various levels (Gray et al. 2015). Overall, many taxa simply do not have published information on their associated traits (Compson et al. 2018, Laigle et al. 2018), or the available publications are geographically biased (Schneider et al. 2019, Keller et al. 2022). Therefore, filling the missing gaps remains a great challenge. Moreover, trait-matching models can be prone to collinearity in traits (e.g., body mass can be correlated with body size; Pichler et al. 2020), to being overreliant on existing trait and interaction databases (Bartomeus et al. 2016, Poisot et al. 2021), and to being biased toward published dietary information (Parravicini et al. 2020), often focusing on the most abundant species. The available data on traits and interactions often do not allow us predict interactions across spatiotemporal scales or to differentiate developmental stages or intraspecific variation (Gray et al. 2015, Bonada and Dolédec 2018, Laigle et al. 2018, Wilkes et al. 2020, Green et al. 2022, Gonçalves-Souza et al. 2023). These data sets also lack complementary predictors of interaction probability, such as relative abundance (Araújo and Rozenfeld 2014, Brousseau et al. 2018a), which provides more information for building food webs than co-occurrences on the basis of presence or absence data (Cazelles et al. 2016, Blanchet et al. 2020).

Reconstructing food webs using eDNA metabarcoding and trait matching

To infer food webs using both eDNA and trait matching, we propose a framework (figure 3) involving the coupling of community structure (taxa × site matrix) and traits (taxa × traits matrix) data sets (figure 3a). The traits data set represents key traits for foraging and vulnerability strategies (e.g., feeding guild, movement type, habitat type; Gravel et al. 2016), which can be compiled from measurements (e.g., body size), literature (e.g., peer-reviewed papers, reports, public databases), inferences from taxonomically similar taxa, field observations, or expert knowledge. Forbidden interactions can be easily identified and pruned (figure 3b) with a set of simple rules (e.g., producers do not consume other taxa). This step limits the scope of inference to possible interactions alone (Morales-Castilla et al. 2015) and increases the efficiency of the following in-depth analysis of trait matching. The proposed framework also requires an interactions data set (taxa × presence or absence of interactions matrix; figure 3c), consisting of known interactions from published studies (e.g., field

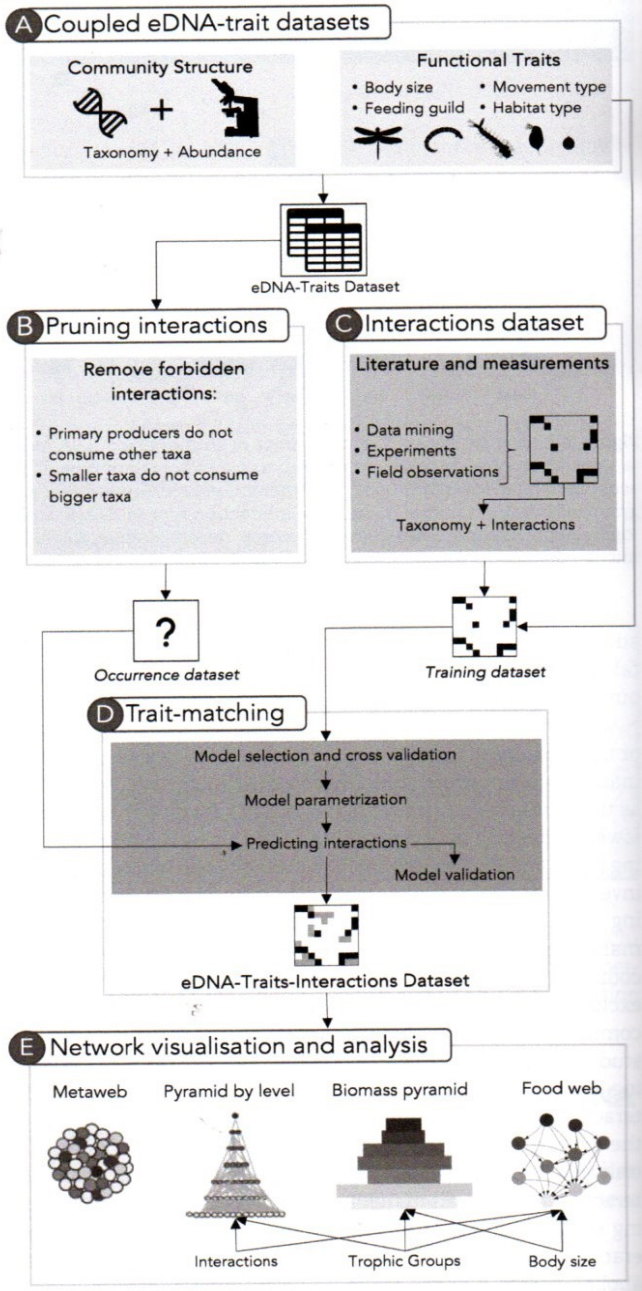


Figure 3. From eDNA to food webs methodological framework. (a) Coupled eDNA–trait data sets are prepared by coupling eDNA (presence–absence) and morphological (abundance) data with traits data sets containing functional traits. (b) Pruning interactions to remove forbidden interactions that are easily identified, trimming the number of possible interactions. (c) Interactions data set, an independent food web database, gathered from published studies. (d) Trait matching, for predicting species interactions, the statistical trait matching is applied to the occurrence data set using the training data set following multiple machine learning models (e.g., random forest, boosted regression trees, ***k***-nearest-neighbor) together with a set of model selection and cross-validations. (e) Network visualization and analysis, the resulting predicted interactions data set is visualized by different means, such as metawebs, food pyramids, and networks.

observations, experiments, and data mining). The trait-matching procedure (figure 3d) requires two types of data sets: training and occurrence data sets. The trait-matching model is initially trained using a data set that includes the known variables of interest (i.e., interactions) and a set of known predictors (i.e., species' traits) of those variables. The training data set is the combination of the interactions and traits for all interacting species in an ecosystem of interest. Following a model selection procedure, to identify the most appropriate statistical model to fit the data, the training data set is used to estimate model parameters that best fit the known interaction probabilities. The occurrence data set consists of unbiased observations of occurrences (i.e., taxa present in samples) and the same trait information as used in the training data set for all observed taxa. The model parametrized using the training data set is then used to predict interaction probabilities on the basis of traits from the occurrence data set (figure 3d). The training model quality depends critically on the resolution of the interaction and the trait data. The resulting matrix of interaction probabilities can then be used to visualize and analyze the resulting networks (figure 3e). Depending on the availability of particular traits, the ecological network information can be illustrated using various approaches ranging from food pyramids to networks. For instance, if traits are not available at a finer resolution, food pyramids striated by trophic level could inform about the general patterns of the network using certain trait databases or published information. However, it should be noted that these food webs cannot accurately determine relationships along gradients of time and space, because they do not infer intraspecific trait variability (see the "Challenges in inferring food webs with trait matching" section).

Although we advocate for the collection of new trait data, it is important to note that our framework can also be used effectively with existing and available trait data sets. However, it is crucial to recognize that the interpretative scope of the resulting insights will always depend on the resolution of the data sets used, encompassing factors such as spatial resolution and intraspecific variation (Bonada and Dolédec 2018, Wilkes et al. 2020, Green et al. 2022, Gonçalves-Souza et al. 2023). Our literature review unveiled the widespread application of both eDNA and trait-matching techniques across diverse ecosystems and taxonomic groups (figure 2). This signifies the growing interest and immense potential of reconstructing ecological networks using taxonomic data generated by eDNA metabarcoding in conjunction with trait-matching models. Our framework is not confined to a specific system; it can accommodate different types of trait and interaction data and different trait-matching models. This nonspecificity enhances its applicability across different research contexts, ecosystems, and taxonomic groups, thereby promoting the advancement of our understanding of ecological interactions at broader scales. Furthermore, the observed variations across food webs result from the distribution of traits in different systems and differences in the underlying rules driving species interactions (Gravel et al. 2016). These trait-matching rules exhibit a degree of generality even across distinct environments (https://doi.org/10.32942/X29K55 [preprint: not peer reviewed]), showcasing the potential applicability of models developed for inferring species interactions in one environment or region to other systems. Ongoing research on model transferability has already demonstrated the reliable predictive ability of trait-based models to infer predator–prey interactions across various ecosystems (https://doi.org/10.32942/X29K55 [preprint: not peer reviewed]).

The Iberian Pond Network case study

We showcase the application of the proposed framework using the Iberian Pond Network (IPN; www.iberianponds.uevora.pt), where trophic interactions were inferred using trait matching (following Laigle et al. 2018), coupling eDNA and trait data (supplemental figure S2 and table S1 in supplemental file S2). The IPN is a multiregion experimental facility of freshwater pond mesocosms distributed across a temperature gradient in the Iberian Peninsula, from southern semiarid to temperate and alpine environments (for details, see supplemental file S2). The IPN final community structure data set (figure 3a) represents a consensus of eDNA- and morphology-based identification. Taxonomic identification was corrected using the lowest level of identification obtained with either approach, and abundances and biomasses, obtained with morphology-based approaches, were redistributed through the eDNA taxa proportionately with their number of reads (Pereira et al. 2021). The resulting data set presented greater diversity and taxonomic resolution with abundance and biomass for 85% of the taxa present. The traits data set consisted of five species' traits (body size, feeding guild, movement and habitat type, and phylogeny) compiled from direct measurements and observations, literature, inferences from taxonomically similar taxa, or expert knowledge (supplemental file S2 table S1).

The IPN metaweb (i.e., the network of all possible trophic interactions within the IPN) included 529 taxa from different trophic levels, with a total of 279,841 potential links (figure 4a). Forbidden interactions were pruned on the basis of major feeding guilds (e.g., producers do not consume animals; figure 4b), removing 72% of all possible links. The predictions from the trait-matching model were trained with the GATEWay database (Brose 2018), an independent interactions data set that included 290 food webs (222,151 feeding links among 5736 species) from five ecosystem types distributed across the globe (Brose 2018, Brose et al. 2019). Only observations from freshwater ecosystems (lakes and streams) were used, including the presence or absence of interactions ($n = 15{,}654$) of 610 taxa and their functional traits (body mass, movement type, metabolic type; training data set), collected from 31 studies (for further details, see Brose et al. 2019). Other trait information that can reflect trophic interactions was added to the functional trait data set, including feeding guild, habitat type, and taxonomy, which was used as a latent trait. Taxonomy was obtained using "taxize" package (Chamberlain and Szocs 2013), and the taxonomic distances between all species were estimated using principal coordinates analysis (first two axes) using the "ape" package (Paradis and Schliep 2019). The same trait information was collected for the IPN data set (occurrence data set; table S1). After model parameterization with the training data set, interactions were further pruned in the occurrence data set and resulted in IPN food web consisting of 518 taxa and 36,913 links (removing 87% of all links), which can be visualized using body mass and total abundance information available as food pyramid and body mass–abundance scaling, respectively (figure 4b). Whereas food pyramids, with trophic level or body mass information, represent an overview of the food web structure, the addition of abundance allows the visualization of body mass–abundance relationships, which provides a significant understanding of community dynamics (Trebilco et al. 2013). Along the IPN temperature gradient, regional food webs were obtained by filtering the interactions of taxa that do not co-occur in the same experimental location (figure 4c). The analysis of these inferred food webs allows

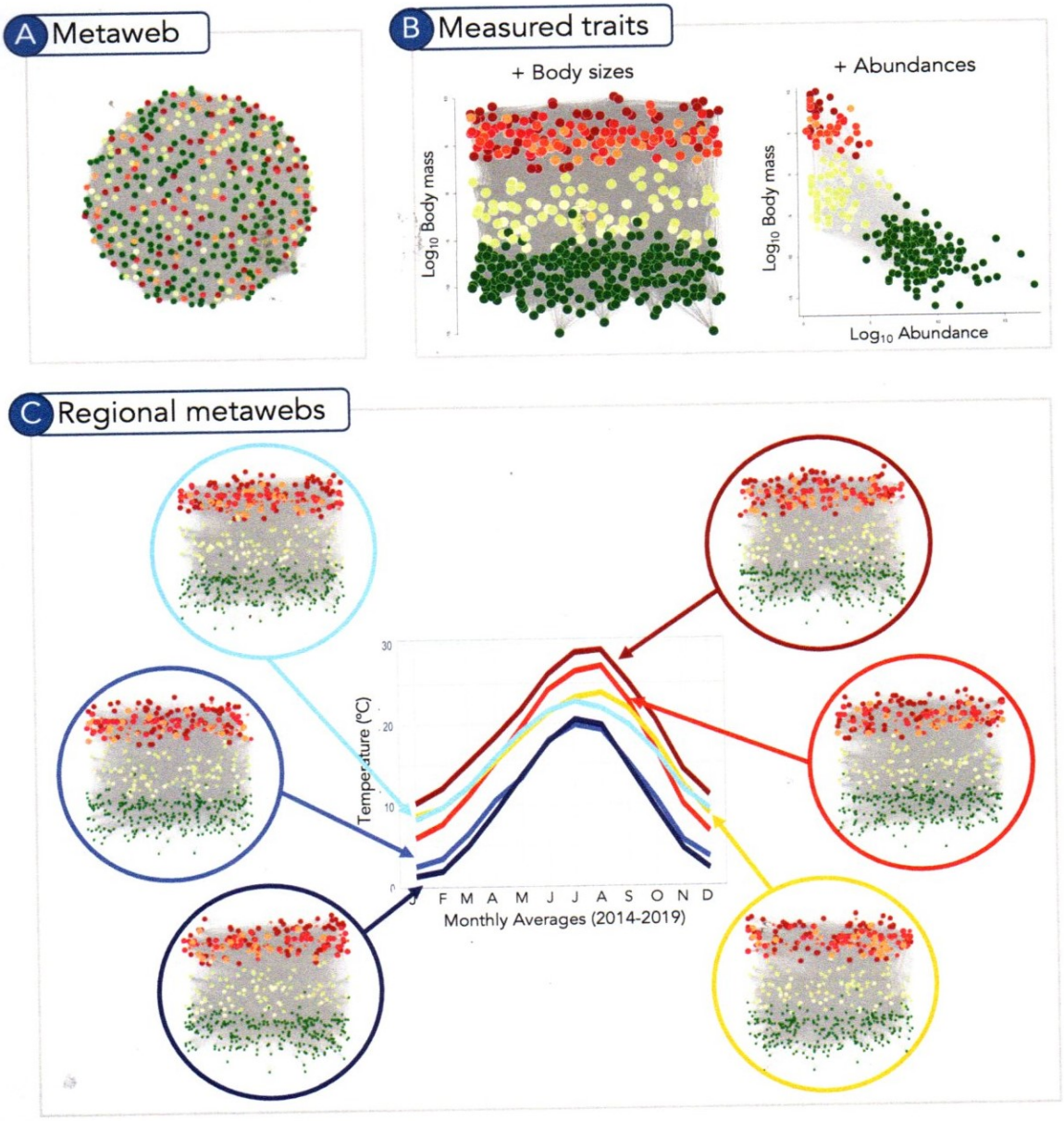


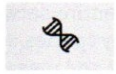
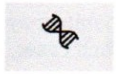
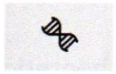
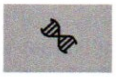
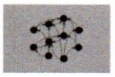
Figure 4. Reconstructing food webs with coupled eDNA-traits recording scheme from the Iberian Pond Network (IPN). (a) IPN metaweb with all possible interactions (529 taxa and a total of 279,841 links). (b) IPN final food web after applying the trait-matching framework (87% of the links were removed). Food web is visualized using body mass (left), for example, to divide taxa into different trophic niches and abundance (right) to understand body mass–abundance scaling patterns. (c) Regional food webs of six experimental locations distributed across a temperature gradient in the Iberian Peninsula. Taxa with no matching traits were not plotted. The colors of the metawebs' nodes indicate different trophic groups: producers, dark green; herbivore zooplankton, light green; omnivore zooplankton, yellow; herbivore macroinvertebrates, orange; omnivore macroinvertebrates, red; and predator macroinvertebrates, dark red. The details are described in supplemental file S2.

us to describe the variation in the trophic networks across environmental and biogeographical gradients.

Toward coupled eDNA–trait recording schemes

eDNA is already being used to construct co-occurrence networks (*sensu* Araújo et al. 2011), providing cost-efficient and rapid analyses of distribution data at larger spatiotemporal scales (table 1; Djurhuus et al. 2020, Seymour et al. 2020). Co-occurrences alone are not evidence of interactions (Cazelles et al. 2016, Blanchet et al. 2020), but they constrain potential interactions (Araújo et al. 2011, Cazelles et al. 2016, Gravel et al. 2019). The reconstruction of food webs from eDNA data sets would benefit from background knowledge—for example, about species traits or existing interactions (Vacher et al. 2016, Compson et al. 2019). Despite ongoing efforts to gather information about observed interactions using text mining from literature and published trait information (Compson et al. 2018, 2019, Danet et al. 2021) or through inferences from interaction data sets (Bloor et al. 2021), there are very few examples of studies reconstructing food webs using trait matching paired with eDNA (Compson et al. 2019, Djurhuus et al. 2020, D'Alessandro and Mariani 2021). Coupled with trait-based

Table 1. Overview of studies using eDNA, morphological, and trait-based approaches.

eDNA Occurrences	Databases		Morphological Occurrences	Measured		Inferences		Outputs		
	Traits	Interactions		Traits	Interactions	Type	Grain	Intraspecific variation	Opportunities	Ref
						Co-occurrence	Taxa	No	Cost-efficient Rapid Finer spatio-temporal scales	a,b
						Co-occurrence	Functional groups	No	Cost-efficient Rapid Finer spatio-temporal scales	c
						Rule-based trait-matching	Taxa Functional groups	No	Cost-efficient Rapid Realistic Finer spatio-temporal scales	d,e
						Statistical trait-matching	Taxa Functional groups	No	Realistic Finer interactions resolution	f,g
						Rule-based trait-matching	Taxa Functional groups	Yes	Realistic Finer interactions resolution	h
						Statistical trait-matching	Taxa Functional groups	Yes	Rapid Realistic Finer spatio-temporal scales Finer interactions resolution	NA

[a]Seymour et al. 2020. [b]Djurhuus et al. 2020. [c]Blackman et al. 2022. [d]Compson et al. 2019. [e]D'Alessandro and Mariani 2021. [f]Laigle et al. 2018. [g]Pecuchet et al. 2020. [h]Pomeranz et al. 2019.

approaches and ideally complemented by traditional taxonomy and metabarcoding to make the link between trait and sequence, eDNA-based food webs can facilitate inferences about species interactions (Compson et al. 2019, D'Alessandro and Mariani 2021) at the interspecific level by using general traits from public trait databases, although they cannot address intraspecific trait variation at higher resolution. Trait matching with morphological data can be used to better estimate the intraspecific variation and potential species interactions, leading to highly resolved food webs (Laigle et al. 2018, Pomeranz et al. 2019, Pecuchet et al. 2020, Pichler et al. 2020). For instance, trait-based modelling approaches were able to predict more than 80% of observed consumer–resource interactions using a set of functional traits, such as body size and feeding guild (Laigle et al. 2018, Pecuchet et al. 2020), despite the coarse resolution of the trait data. However, because of great efforts associated with taxonomy and trait measurements, trait-matching approaches with morphological data fail to provide rapid and cost-efficient responses. Therefore, eDNA-based recording schemes will greatly benefit from a coupled collection of morphology-based trait data, through well-designed recording schemes (e.g., random or stratified sampling), which bring finer scale and realistic inferences of food webs with greater inter- and intraspecific resolution (table 1). We propose the development of eDNA–trait recording schemes, which should include morphological samples at relevant spatial and temporal scales to ensure that eDNA surveys will have suitable background information for posterior food-web analysis. A sampling strategy targeting food-web responses should aspire to maximize the coverage quality of eDNA and intraspecific trait measurements across relevant gradients and scales. We have identified key elements that are worth consideration:

Identify and measure relevant traits

Ideally, eDNA surveys should be coupled with new trait measurements that are relevant to the ecological mechanisms driving the responses being investigated. The selection and availability of trait information are critical to inferring food webs. The sensitivity of trait matching to the selection of traits and collinearity in traits' measurements (i.e., redundant traits) can influence the ability to infer interactions (Pichler et al. 2020). Some traits (e.g., body size, phylogeny) might be better predictors of trophic interactions than others (e.g., movement and habitat types, prey capture strategy; Gravel et al. 2013, Gray et al. 2015, Brousseau et al. 2018a, Laigle et al. 2018, Pomeranz et al. 2019). Body size is a good predictor of interactions across multiple ecosystems and taxonomic groups (Brose et al. 2019), but this might not be true for complex ecosystems with indirect and intraguild interactions (Jonsson et al. 2018). If body size is being measured and considering the correlation between body size and other morphometric traits (Rogell et al. 2020), it might be advantageous to focus on other traits (e.g., phylogeny, Pomeranz et al. 2019; habitat use, Albouy et al. 2019; foraging mode, Wootton et al. 2022). Such prioritization should be aligned with the hypothesized mechanisms driving interactions (Brousseau et al. 2018a). Because phylogenetically similar taxa share multiple traits, which could contribute to regulating trophic interactions (Eklöf et al. 2012), genealogical relationships among species have been shown to be a predictor of interspecific interactions under the assumption that key biological traits are conserved (e.g., Eklöf et al. 2013, Romdal et al. 2013, Bennett et al. 2021). The macro- and microhabitat preferences of organisms could also change co-occurrences and, therefore, interaction probabilities, thereby adding more resolution to the inferred food webs (Morales-Castilla et al. 2015).

Quantify intraspecific trait variation

Intraspecific trait variation has been reported for a range of systems and taxonomic groups (Novak and Tinker 2015, Jackson et al. 2017, Herrando-Pérez et al. 2020), and its impacts on predictions of individual species distributions are potentially great (Valladares et al. 2014). Given that changes in intraspecific trait variation are not detected using eDNA approaches, the ability of this approach to predict changes in ecological interactions across space and time can be impaired (González-Varo and Traveset 2016, Pellissier et al. 2018). Therefore, eDNA data

sets would benefit from inter- and intraspecific trait variation data obtained with morphology-based approaches. For example, in Pereira and colleagues (2021), eDNA metabarcoding and morphology-based data sets were paired on the basis of a framework with increasingly restrictive workflows, following a series of taxonomic and geographical filters. This framework allowed taxonomic resolution to be improved for 30% of the taxa and enabled the species' traits (e.g., body size) with intraspecific variation and abundance to be assigned to 85% of the taxa in the final data set (Pereira et al. 2021).

Incorporate relative abundances and ontogeny as predictors

The interaction strength between species pairs can be estimated using relative species abundances (Araújo and Rozenfeld 2014, Cazelles et al. 2016), a prediction that has been verified in food webs (Canard et al. 2012), host–parasite bipartite networks (Canard et al. 2014), and mutualistic plant–animal networks (Donoso et al. 2017). According to the neutral theory, the probability of observing a pairwise interaction is dependent on the abundance of the interacting species (Vázquez et al. 2009), implying that the more abundant the species are, the stronger will be their pairwise interaction (Araújo and Rozenfeld 2014, Cazelles et al. 2016). Rare species tend to be more specialized in how they interact, simply because they have fewer opportunities to establish the diversity of potential interactions (Henriksen et al. 2022). Stratified sampling across relevant environmental gradients would facilitate the assignment of interaction probabilities on the basis of the abundances of the surveyed species (Pereira et al. 2021). An analogous example is ontogeny, species often assume distinct functional roles throughout their life history stages, resulting in important effects on interaction probabilities (Brousseau et al. 2018b) and therefore on food-web structure but cannot be detected directly using eDNA because of its inherent limitations. Collecting targeted measurements of life-history events, which can vary across seasons (e.g., developmental stage) and temporal dynamics (e.g., response to impact, emergence timing), together with eDNA would allow capturing such differences and incorporating them into food-web inferences via trait matching (Olivier et al. 2019, Schleuning et al. 2020).

Pair diet analysis with eDNA–trait recording schemes

DNA metabarcoding, with high-throughput sequencing technologies, has already been embraced by studies on diet analysis (i.e., determination of feeding habits and resource preferences; for reviews, see Clare 2014, Alberdi et al. 2019). However, DNA-based diet analysis has intrinsic limitations that will always be a challenge. For example, DNA metabarcoding alone cannot discriminate secondary consumption (taxa present in the stomach of preyed organisms) and cannot detect cannibalism, because consumer and resource sequences cannot be disentangled (Ray et al. 2016, Siegenthaler et al. 2019). Food webs inferred from eDNA and trait matching provide both visual and quantitative information on food-web structure. Complementing it with DNA metabarcoding or other diet approaches, such as direct observations of foraging behavior, sorting and identification of resource fragments from stomach or fecal content, or tracking stable isotopic markers, can clarify the trophic niche widths, energy pathways, and community functioning and resource use (Compson et al. 2019), allowing for comparative studies of trophic networks across space and time. Regarding trophic and diet-related traits, stable isotope analysis, for example, can provide valuable estimates of inter- and intraspecific trait variation (Bearhop et al. 2004), which can be incorporated into food-web reconstruction with trait matching modelling to refine depiction of species' diets. Targeting highly connected species critical to food webs, prior to sampling the entire network (Compson et al. 2019), allows the identification of key species, such as large consumers, which are often easier to collect and measure. Furthermore, individual specialization of feeding habits may occur within some species (Layman et al. 2015); in such cases, targeted sampling might be particularly relevant.

Close integration with modelling and theoretical predictions

Although there has been an increase in species occurrence data, high-quality empirical data on ecological interactions remain challenging (Gravel et al. 2019, Bennett et al. 2021, Mestre et al. 2022a). The lack of comprehensive data on species interactions, which is often aggregated at high taxonomic levels, leads to a loss of critical information to distinguish between individual-level and whole-species interactions (Wells and O'Hara 2013, Strydom et al. 2021), potentially introducing biases into understanding of food-web structure (De Aguiar et al. 2019). To predict species interactions and to describe the food-web structure across environmental gradients, one must build on existing incomplete data (Strydom et al. 2021), experiments, observations, and modelling to enable the formulation of quantitative predictions that can be rigorously tested (Mestre et al. 2022b). Recent advances in food-web research have been focused on mapping and understanding variations in food-web structures from continental to global scales (Albouy et al. 2019, Mendoza and Araújo 2019, O'Connor et al. 2020, 2022) and on generating testable hypotheses (Baiser et al. 2019). To fill this data gap and improve the understanding of food webs, it is critical to advance theoretical frameworks, develop models, and gain deeper insights into the realized versus potential food-web dichotomy (Morales-Castilla et al. 2015, Mestre et al. 2022a). Moreover, the inference of interactions can be used to target field work by addressing the data challenge associated with the exponential increase in potential interactions with increasing species richness. However, the number of links does not increase in the same way; consequently, one could prioritize field work only on predicted interactions, resulting in significant savings in time and effort.

Optimizing eDNA–trait recording schemes

Public databases (e.g., GenBank) are reliable resources (Lang et al. 2019), especially at taxonomic levels above species (Locatelli et al. 2020), but there is still a need for more comprehensive sequence databases, especially for understudied regions and trophic groups (Ruppert et al. 2019, Weigand et al. 2019). For example, although more than 80% of all fish species are referenced in public databases, only 26% of the marine invertebrates used in aquatic biomonitoring in Europe are covered (Weigand et al. 2019); this percentage can be even lower in less explored regions or with higher overall diversity. More effective quality assurance and control of barcode reference libraries are required (for standards and guidelines for species and DNA curation of reference databases, see Rimet et al. 2021). eDNA identifications are most exact when they are based on comprehensive DNA reference libraries underpinned by robust taxonomic frameworks and physical voucher specimens (Taberlet et al. 2012, Berry et al. 2021), which are rarely obtained, often unidentifiable, or unavailable. The idea of designing coupled eDNA–trait recording schemes

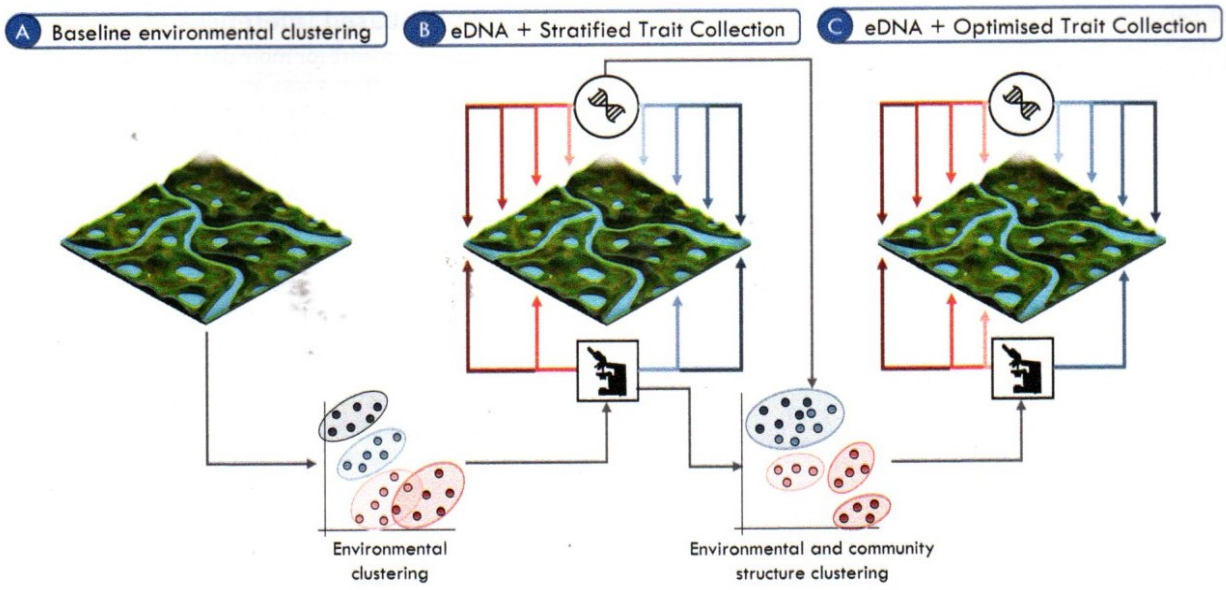


Figure 5. Heuristic stratification and optimization of trait collection based on environmental and community structure clustering. (a) Baseline environmental clustering. (b) Coupled eDNA and stratified trait collection based on environmental clustering. (c) Coupled eDNA and optimized stratification of trait collection based on environmental and community clustering. In panels (b) and (c), trait information is collected at relevant spatial and temporal scales to ensure eDNA samples have appropriate background information for food-web reconstruction. Recording and monitoring schemes should aim to not only maximize eDNA coverage of target ecosystems, but also the quality of reference libraries, traits, and interactions data. The red and blue arrows indicate possible sampling sites across any spatial or environmental gradient (e.g., temperature, salinization, urbanization).

goes against some of the original premises of eDNA studies (e.g., Taberlet et al. 2012), which were to maximize the cost per sample and reduce the total cost of biodiversity recording schemes, but will provide an added resolution and future value to eDNA data sets.

A possibility to circumvent these handicaps is to devise hypothesis-driven stratified recording strategies that account for the scale or relevant environmental gradient that is likely to drive the responses of interest (figure 5). Stratification (or optimization) of trait measurements consists of identifying relevant partitions across temporal (e.g., seasons) or environmental gradients (Hortal et al. 2009) that reflect the trait's intraspecific variation across such gradients while not rendering the entire recording and monitoring schemes unsustainable. Stratification of trait collection can be based on environmental clustering (e.g., principal component analysis, fuzzy clustering using machine learning; figure 5a; Güler et al. 2012, Mendoza and Araújo 2019, 2022) of sampling sites along environmental gradients (e.g., temperature, salinization; Kortsch et al. 2019) where trait collection follows the obtained clusters (figure 5b). If data on community structure are available, an optimization of the stratified recording schemes can be made (figure 5c) following environmental and community structure clustering of sampling sites. When measuring the community response to disturbance (e.g., land-use change; Castro et al. 2018), it should be ensured that coupled eDNA–trait samples are taken from affected and control areas. Potential changes in the food-web structure between warmer and colder or affected and control areas (e.g., changes in body sizes, phenology) could pass undetected if using only eDNA and if the target trait had no matching spatial resolution (i.e., collected from the literature and trait databases) to account for trait variation. Stratified trait collections allow us to establish differences in key traits and incorporate those in estimates of changes in the food-web structure. Similarly, if recording and monitoring schemes compare locations across large spatial and temporal scales (e.g., different basins, regions, or islands across seasons and years) using previously collected trait data to reconstruct food webs, they are unlikely to capture food-web level changes, because traits sourced from literature are often reported as species-level averages (Gallagher et al. 2020). Coupled eDNA–trait recording schemes involve greater labor and costs, but the stratification and optimization of traits collection may reduce the load and, if they are collected across relevant geographic or environmental and temporal gradients, it would maximize the likelihood of detecting changes in food webs. By doing such optimization, labor and costs would increase by 1.4–1.9 times more than doing eDNA alone (estimates from the IPN case study, which may vary across countries and sampling design).

Future developments

We highlight four areas of development that include the most urgent gaps to be addressed when resolving food webs from eDNA–trait recording schemes.

Fine-tuning eDNA approaches to resolve food webs

With the fast development of eDNA and its application to a wide range of research and systems, a plethora of methodological variations was introduced at all stages of the workflow (Seymour 2019). There is a remarkable degree of consensus on best scientific practices, and recent and ongoing efforts are being made to provide guidelines (e.g., EU COST Action DNAqua-Net; Bruce et al. 2021) for every step of the field and laboratory workflows. However, this may not be readily discerned from the now extensive body of research literature, and furthermore, research is required in some areas (e.g., bioinformatics; Walters et al. 2019) to

balance best practices with the constraints and priorities of end users (Bruce et al. 2021). Increased taxonomic resolution of all target biodiversity is needed to resolve food webs at finer levels (i.e., higher than trophospecies). To provide better overall coverage at the species level, the design and optimization of primers (both specificity and generality; Derocles et al. 2018) are paramount or, to completely avoid primer biases, moving toward shotgun sequencing (comprehensive sequencing of all genes from organisms present in a sample without relying on amplification by polymerase chain reaction [PCR]; Ruppert et al. 2019) or hybridization capture (a target enrichment method for obtaining DNA data from samples with high nontarget content; Sigsgaard et al. 2020). PCR amplification is still predominantly used in eDNA studies because of the high costs associated with achieving the sequencing depth required for shotgun sequencing (Cuff et al. 2022) and other associated challenges (e.g., the lack of sufficiently curated reference genomes; Parducci et al. 2019).

Except for a few well-studied taxa (e.g., birds and fishes), the number of species that actually exist on Earth is still uncertain (Mora et al. 2011, Scheffers et al. 2012, Costello et al. 2013). Furthermore, only a fraction of those species catalogued to date are represented in sequence reference databases. Despite ongoing efforts to improve the accuracy of taxonomic identification of eDNA data by filling the gaps in barcoding reference databases (Albaina et al. 2016, McGee et al. 2019, Weigand et al. 2019), there is a critical need to both make the content of such databases truly representative of life on Earth and pair genetic reference data with curated voucher specimens (Taberlet et al. 2012, Berry et al. 2021). It is essential to create advanced facilities (Ruppert et al. 2019) to house the growing data and reference libraries, through artificial intelligence, text mining pipelines, and open-source and cloud-based tools. In this way, we can ensure the use of these data sets to provide more complex information to citizens and decision-makers for management at local and regional scales (Makiola et al. 2020). There is no doubt that the value of eDNA would be significantly boosted by incorporating sequence data and respective taxonomy into not only nucleotide reference databases (e.g., NCBI [the National Center for Biotechnology Information]) but also into more accessible and user-friendly databases, such as biodiversity portals (e.g., Aquatic eDNAtlas; GBIF [the Global Biodiversity Information Facility]). Such a shift could centralize eDNA biodiversity data where it could be integrated with global interaction databases (e.g., GloBI, Mangal, Web of Life). This integration will be improved by a parallel collection of molecular and morphological data in a subset of reference points along larger temporal scales, being also essential to assure backward and forward compatibility of time series data sets.

Reconstructing food webs using only eDNA data faces a major challenge, which is the lack of information on species abundance and biomass. Although it is promising, the positive correlation observed between DNA concentration and the abundance of focal species (e.g., fish; Rourke et al. 2021) might not directly extend to multitrophic communities. Moreover, the variance in eDNA concentrations (or read sequence numbers) and species abundance across studies in artificial and natural systems adds further complexity (Yates et al. 2019). In this context, the use of density or biomass has emerged as more reliable indicators of read counts (e.g., Li et al. 2021). Further research is warranted to fully comprehend these relationships; however, their initial implications do appear promising for food-web studies using eDNA.

Improve trait-based inferences of food webs

Despite the requirements for more data, trait matching using machine learning algorithms was successful at inferring food web structure in several studies (Laigle et al. 2018, Mendoza and Araújo 2019, Pomeranz et al. 2019, 2022). There are still issues to resolve, such as data availability, data resolution and quality, and methods related to inferring interactions. Real-time and automated biomonitoring tools for environmental data recording and processing are already available, and novel frameworks are underway to count, track, and even record certain behavioral, functional, and morphological traits (for a review, see Besson et al. 2022). Such advancements will open new opportunities for producing multidimensional and high-resolution data to refine trait-based predictions of biotic interactions. In addition, recent studies proposed automating trait collection from the literature (Compson et al. 2018, 2019). Efforts should be made in developing and adopting systematic protocols for the collection of trait measurements that are specific and that target trait-matching approaches (Moretti et al. 2017). There is also a need to validate trait-matching workflows using known interaction databases (e.g., GloBI, Mangal, Web of Life) to assess the sensitivity of different resolutions in available trait data. To efficiently predict the impacts of species range shifts on food webs, trait-based approaches should be merged with ecological network analysis (Pecuchet et al. 2020), and other drivers of interactions should be integrated (e.g., co-occurrence, relative abundance; Brousseau et al. 2018a). Finally, to facilitate the mapping of trophic interactions at a global level (e.g., Mestre et al. 2022b), efforts should be made to unify open global databases, new computing tools (e.g., machine learning; Pichler et al. 2020, Pichler and Hartig 2023), and statistical approaches (Albouy et al. 2019), through data standardization and harmonization and by filling gaps for rare taxa and taxa from globally underrepresented regions (Poisot et al. 2021, Keller et al. 2022).

Apply trait-matching approaches to ancient sedaDNA data sets

Advances in the extraction and identification of DNA that was shed into the environment and preserved over time in, for example, lake sediments and permafrost (so-called sedaDNA), provide a way to augment discontinuous paleontological assemblages (Domaizon et al. 2017) and to reconstruct past conditions of these ecosystems (Ruppert et al. 2019). When the cocollection of morphological data is not possible, reconstructions of past food webs can be achieved by combining morphological analyses of organisms retained in the sediment record (e.g., diatoms, cladocerans, insect larvae; Raposeiro et al. 2021) and inferences of traits based on modern trait distributions for species only detected by sedaDNA (Fordham et al. 2016, Nogués-Bravo et al. 2018). These approximations of the overall structures of past ecosystems and how they changed over time could be used as baselines for future conservation planning, including backcasting testing of climate change models (Sønstebø et al. 2010), modeling invasive species emergence (Ficetola et al. 2018), and the assessment of anthropogenic influences on past biodiversity and landscapes (Raposeiro et al. 2021). However, because each ancient sediment sample can integrate DNA shed across multiple years (depending on DNA degradation and sedimentation rates), the temporal resolution will depend on the width of the time period covered by the sample (Anderson-Carpenter et al. 2011) and the completeness of the DNA record (Giguet-Covex et al.

2019). Broader integration of sediment layers will inevitably result in more general approximations of the ecosystem structure, which will be suitable for detecting long-term transitions between ecosystem states rather than measuring short-term responses (e.g., seasonal). Further steps are needed in the expanding world of DNA-based approaches to address some constraints associated with sedaDNA data sets, including understanding the preservation mechanism of DNA molecules in ancient sediments and extending the time window of this technique in the future (Jia et al. 2022).

Incorporating environmental RNA to derive functional information

Similar to eDNA, RNA is also shed by organisms and found in environmental samples, such as water, soil, sediment, and air, and is referred to as *environmental RNA* (eRNA; Cristescu 2019). Because RNA is expressed by physiologically active organisms and has a fast turnover rate, it has been argued that it can provide more accurate insights (and finer spatiotemporal inferences) into the diversity of living species in an ecosystem (Cristescu 2019, Marshall et al. 2021, Yates et al. 2021, Giroux et al. 2022). Furthermore, another potential advantage of eRNA over eDNA is argued to be that RNA has the potential to provide functional information about the state of a given community or population and, therefore, the ecosystem as a whole (Veilleux et al. 2021, Giroux et al. 2022). However, although the successful characterization of RNA shed by macroorganisms into their environment has been demonstrated (e.g., Marshall et al. 2021, Giroux et al. 2022, Littlefair et al. 2022), the profiling of environmental transcriptomes (i.e., gene expression profiling based on eRNA) remains at the proof-of-concept level (Veilleux et al. 2021). In this regard, priority areas for future research on eRNA have been identified (Yates et al. 2021): understanding eRNA production, because gene expression is tissue dependent (Stevens and Parsley 2023); understanding eRNA degradation, especially the differential degradation rates of different forms of RNA (Cristescu 2019); and improving sample preservation to improve the quantity and quality of RNA collected and preserved in environmental samples (Yates et al. 2021). The resolution of eRNA-related challenges could potentially position it as a more powerful tool than eDNA for assessing community responses to environmental changes across trophic levels (Cristescu 2019, Giroux et al. 2022).

Move from binary networks to energy flow models

Energy flow models can be used to estimate transfer efficiency using information on trophic levels and production at each level (Eddy et al. 2021). Food webs inferred with trait matching provide an approximation of the main energy pathways but may lack details needed to resolve species-specific energy pathways (Brown and Gillooly 2003). Recent methodological and analytical developments have revealed the potential of using eDNA samples for estimating catchment-wide biomass distributions (Carraro et al. 2018, Shelton et al. 2019), which are valuable for inferring interactions (Brose et al. 2019). Coupling inferences from food-web structures using trait matching and measures of production at each trophic level can quantify the relationships between energy flows and biodiversity (Eddy et al. 2021), potentially unlocking the power of eDNA for monitoring increasingly complex ecosystem dynamics across environmental gradients.

Conclusions

On the basis of a critical review of the literature, a case study, and a synthesis of challenges and opportunities emerging from it, we propose a framework to combine eDNA and trait data with trait-matching algorithms to describe and detect changes in food webs across environmental gradients. We advocate the need for future-proofing biodiversity data sets on the basis of eDNA approaches, anticipating their potential use in detecting past and future responses to change. Despite increasing the cost per sample, biodiversity recording schemes that couple eDNA with stratified trait collection would advance the development of food web studies using eDNA in the future by increasing the food-web resolution at both spatiotemporal and trait-based scales. At a time of great development and diversity of eDNA applications, beyond ensuring the molecular aspects of these approaches meet sampling, laboratory, and bioinformatics standards, we must make sure that traits are also collected at scales that are relevant to the questions being asked. For example, when studying the consequences of phenological shifts, as a result of climate change, on species interactions and food-web structure, phenological traits should be sampled across seasons and years. eDNA data sets can be used together with trait databases and measurements, not only for biomonitoring but also to explore the trophic structure of communities and may support management interventions (D'Alessandro and Mariani 2021).

Supplemental Material

Supplemental documents are available at *BIOSCI* online.

Data and scripts for data analysis and figures are available on the GitHub repository https://github.com/IberianPonds/manuscript_trait_matching.git.

Acknowledgments

This study was funded by the Portuguese Science and Technology Foundation (FCT) through the scientific projects TrophicResponse—trophic responses to macroecological gradients (project no. PTDC/BIA-BIC/0352/2014), and StateShifts—predicting state shifts in energetic food webs under climate change (project no. PTDC/AAG-MAA/3764/2014) and European Regional Development Fund (FEDER) through the COMPETE 2020 Programme and National Funds through FCT under projects no. POCI-01–0145-FEDER-007688 and no. PTDC/CTA-AMB/30793/2017 (AdaptAlentejo—predicting ecosystem-level responses to climate change). CLP acknowledges support by FCT through PhD studentship (grant no. SFRH/BD/102020/2014). MBA and MGM were funded through AQUACOSM-PLUS (grant no. H2020-INFRAIA-2019–1, project no. 871081). MGM was supported by grants no. RyC-2016–19348 and no. RTI2018-095749-A-I00 funded by grant no. MCIN/AEI/10.13039/501100011033, "ESF Investing in your future" and "ERDF A way of making Europe." We thank Núria Galiana Ibáñez for her comments and suggestions on earlier versions of the manuscript. We also thank three anonymous reviewers for their comments and suggestions.

Author contributions

Cátia Lúcio Pereira (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing), Zeynep