Towards all-inclusive community ecology via DNA metabarcoding

Gentile Francesco Ficetola[1] and Pierre Taberlet[2]

[1]Università degli Studi di Milano
[2]Université Joseph Fourier

October 11, 2022

Abstract

An exhaustive assessment of biodiversity is a major challenge of ecological research, and molecular approaches such as the metabarcoding of environmental DNA are boosting our ability to perform biodiversity inventories. Are we actually able to assess the whole community, to unravel the intricate interactions between organisms and the impacts of global changes on the different trophic levels? The majority of metabarcoding papers published in the last years used just one or two markers and analyzed a limited number of taxonomic groups. Nevertheless, approaches are emerging that might allow "all-taxa biological inventories". Exhaustive biodiversity assessments can be attempted by combining a large number of specific primers, by exploiting the power of universal primers, or by combining specific and universal primers to obtain good information on key taxa while limiting the overlooked biodiversity. Multiplexes of primers and shotgun sequencing may provide a better coverage of biodiversity compared to standard metabarcoding, but still require major methodological advances. We identify the strengths and limitations of different approaches, and suggest new development lines that might improve broad scale biodiversity analyses in the near future.

Towards all-inclusive community ecology via DNA metabarcoding

Gentile Francesco Ficetola[1,2]*, Pierre Taberlet[2,3]

1) Department of Environmental Science and Policy, Università degli Studi di Milano, Via Celoria 26, 20133 Milano, Italy

2) Univ. Grenoble Alpes, CNRS, Univ. Savoie Mont Blanc, LECA, Laboratoire d'Ecologie Alpine, F-38000, Grenoble, France

3) UiT – The Arctic University of Norway, Tromsø Museum, Tromsø, Norway

* Corresponding author. francesco.ficetola@gmail.com

Abstract

An exhaustive assessment of biodiversity is a major challenge of ecological research, and molecular approaches such as the metabarcoding of environmental DNA are boosting our ability to perform biodiversity inventories. Are we actually able to assess the whole community, to unravel the intricate interactions between organisms and the impacts of global changes on the different trophic levels? The majority of metabarcoding papers published in the last years used just one or two markers and analyzed a limited number of taxonomic groups. Nevertheless, approaches are emerging that might allow "all-taxa biological inventories". Exhaustive biodiversity assessments can be attempted by combining a large number of specific primers, by exploiting the power of universal primers, or by combining specific and universal primers to obtain good information on key taxa while limiting the overlooked biodiversity. Multiplexes of primers and shotgun sequencing may provide a better coverage of biodiversity compared to standard metabarcoding, but still require major methodological

advances. We identify the strengths and limitations of different approaches, and suggest new development lines that might improve broad scale biodiversity analyses in the near future.

Key words:

Environmental DNA; freshwater and marine biodiversity; multi-trophic analyses; primer cocktails; shotgun sequencing; soil biodiversity

1. INTRODUCTION

An exhaustive assessment of biodiversity has always been a major challenge for ecologists. In principle, all the organisms living in an ecosystem can interact with each other: some insects and mammals feed on plants, plants interact with soil fungi, protists can feed on bacteria or parasitize other eukaryotes, and of course many other interactions occur. Ideally, we should assess the occurrence (and perhaps the abundance) of all the organisms, shall we want to unravel the impact of environmental changes on biodiversity, eventually taking into account the potential biotic interactions. Unfortunately, we all know that this is only rarely possible. If we want to use traditional approaches (e.g. morphological identification of species), thousands systematists should work together for weeks to produce an "all-taxa biological inventory" of just a hectare of tropical forest (Lawton et al., 1998). The emergence of molecular approaches (starting with DNA barcoding) has certainly revolutionized biodiversity inventories, as it allows a much faster and cheaper assessment of present species, particularly for taxonomic groups including many difficult to identify, cryptic or undescribed taxa (Floyd, Abebe, Papert, & Blaxter, 2002; Hebert, Cywinska, Ball, & DeWaard, 2003; Hebert, Penton, Burns, Janzen, & Hallwachs, 2004). DNA metabarcoding now allows the contemporary assessment of a huge number of species, starting from both environmental samples and from tissues (as nicely shown by many papers in this special issue). Does this mean that we are finally able to assess the whole community, and to unravel the intricate interactions between organisms?

So far, beside rare exceptions, this does not seem the case. Out of 70 papers using DNA metabarcoding to study biodiversity variation published in target journals during the last two years (see supplementary materials for details on methods), the majority used just one or two primer pairs and focused on just one (e.g. arthropods, fish, fungi, plants...) or two taxa (e.g. plants + mammals; bacteria + micro-eukaryotes; Fig. 1a; Supplementary Table S1). Several studies had a broad taxonomic scope and used generalist primers (particularly targeting COI and 18S) to amplify very broad groups (e.g. all the eukaryotes, all the animals...), while very few attempted an exhaustive biodiversity analysis using multiple primer pairs each of which targets a different taxon (Fig. 1).

Nevertheless, in principle several strategies may be adopted to obtain detailed information over a broad spectrum of taxa, and attempt a nearly-complete reconstruction of communities on the basis of DNA metabarcoding related approach that might disclose new avenues to biodiversity and ecological research. In this short contribution, we describe some of these approaches, we discuss their strengths and limitations (Table 1), and suggest new development lines that might improve broad scale biodiversity analyses in the near future.

2. POTENTIAL STRATEGIES FOR ALL-INCLUSIVE BIODIVERSITY ANALYSIS USING MOLECULAR APPROACHES

2.1 Using a large number of metabarcodes in the same study

A very large number of primers has been developed and tested for metabarcoding studies. For instance, Taberlet, Bonin, Zinger, and Coissac (2018) proposed 62 distinct primers pairs for DNA metabarcoding, some of which were extremely generalist and amplified very broad taxa (e.g. all the bacteria and archaea; all the eukaryotes...) and others being much more specific, focusing on well-defined taxa (e.g. turtles, the plant family Asteraceae...). In principle, we can amplify the eDNA extracted from one single environmental sample using multiple primers, and then combine the results to attempt an overall reconstruction of biodiversity (Jurburg, Keil, Singh, & Chase, 2021). For example, we might study soil biodiversity by analyzing markers specific for bacteria, fungi, earthworms, insects, springtails..., while freshwater diversity can be assessed by

combining primers that amplify bacteria, protists, insects, fishes, amphibians... (Guerrieri et al., 2022; Li, Qin, Wang, Zhang, & Yang, 2023).

Combining multiple markers allows a good resolution for the selected focal taxa, particularly if each marker has a well-defined and limited taxonomic scope. The integration of results of different primers can allow assessing the response of multiple taxa to environmental gradients, and even attempting the reconstruction of interaction networks (Li et al., 2023).

Unfortunately, targeting multiple taxa increases the cost and labor associated with the laboratory and sequencing, as using many markers means running many PCR reactions, sequencing lanes and so on. Furthermore, even if unlimited resources were available (and this is rarely the case), the amount of eDNA available for amplification remains limited. Imagine you have extracted 100 µL of eDNA from water, each PCR reaction requires 2 µL of template DNA, and you want to run eight replicated PCRs per sample to detect rare species with a limited rate of false negatives (Ficetola et al., 2015). In this case, the template DNA is only enough for a maximum of six primers, thus some key taxon will always be missed. For instance, if we analyze water biodiversity using primers amplifying bacteria, diatoms, mollusks, insects, fish and amphibians we will miss key taxa such as crustaceans and most of micro-eukaryotes.

2.2 Using very generalist or degenerated primers

In principle, we might choose a few very generalist, universal primers, such as the ones amplifying all the eukaryotes or most of the animals (e.g. 18S rDNA or COI-based primers). Several studies have adopted this approach (Fig. 1b) as it has clear advantages, including relatively cheap cost, and relatively easy implementation (see Jurburg et al., 2021 for additional discussions on limitations and recommendations). In principle, with 2 / 3 primer pairs (e.g. one eukaryote and one prokaryote marker) we might try amplifying the whole tree of life (e.g. Holman et al., 2021; Martinez-Almoyna et al., 2019). Unfortunately, the search for perfect, truly universal primers has been compared to the search for the Holy Grail (Rubinoff, Cameron, & Will, 2006). On the one hand, some "universal" primers have limited resolution, or have heterogeneous resolution across the three of life. For instance, some primer pairs focusing on 18S (e.g. the Euka02 primer pair, Guardiola et al., 2015) amplify most eukaryotes and have a reasonable resolution for some taxonomic groups (e.g. nematodes), but a very poor resolution for other taxa (e.g. plants), with complex consequences for data analyses (Jurburg et al., 2021). On the other hand, generalist primers such as those amplifying COI have heterogeneous amplification rate among the target species. The taxa with less mismatches will be amplified preferentially, and this can reduce the success over other taxa. Highly degenerated primers have additional issues such as frequent amplification of non-target regions, and the amplification of non-target taxa (e.g. bacterial DNA amplified with COI primers) (Hintikka, Carlsson, & Carlsson, 2022).

Recently, long-read metabarcoding has been proposed to overcome the limited resolution of many generalist primers (Jamy et al., 2022). With this approach, a very long (e.g. 4500 bp) DNA fragment is amplified with universal primers and then processed through technologies that allow the sequencing of long reads (Jamy et al., 2022). The long-read metabarcoding provides unprecedented taxonomic resolution compared to traditional generalist primers, still poses major technical issues (e.g. chimaera formation) and is much more expensive than short-read metabarcoding. Furthermore, long-read metabarcoding is still rarely applied, and several aspects of this approach will deserve future adjustments and analyses, including the actual universality of primers.

2.3 Combining very generalist and more specific primers

In order to overcome the limitations of strategies 2.1 and 2.2, it is possible to analyze the same environmental DNA using both specific primers targeting taxa with particular ecological role (e.g. high taxonomic diversity, keystone functions...), and generalist primers. For instance, for the analysis of soil biodiversity we might complement primers amplifying insects, springtails, earthworms and fungi, with a primer that amplifies all the eukaryotes and can give an idea of the diversity for groups not amplified with the previous ones (micro-eukaryotes, nematodes, rotifers...) (Bloor, Si-Moussi, Taberlet, Carrère, & Hedde, 2021; Calderón-Sanou et al., 2022; Guerrieri et al., 2022). This approach has the advantage of providing a reasonable representation

of biodiversity, with good information on selected key taxa and few taxa completely missing, and might thus allow exploring complex relationships between multiple taxonomic groups (Bloor et al., 2021; Calderón-Sanou et al., 2022). Nevertheless, similarly to approach 2.1, it remains costly and labor-intensive.

Furthermore, with this approach, the resolution of markers can be extremely heterogeneous among taxa amplified by specific and generalist primers. For instance, the above-cited combination of primers would result in an excellent taxonomic resolution for earthworms and springtails, but a very coarse one for other taxa (e.g. rotifers). Combining taxonomic tables with very different resolution in ecological analyses can be extremely complex, and comparing the biodiversity (e.g. taxonomic richness) of taxonomic groups amplified with different markers is certainly problematic. Even if some analytical strategies can help combining information from disparate groups (Jurburg et al., 2021), understanding the consequences of analyzing altogether taxa with very different taxonomic resolution remains a major methodological challenge associated with this approach.

2.4 Multiplex of primers

An alternative approach is combining multiple metabarcoding primers in the same PCR mix, to simultaneously amplify and sequence multiple taxonomic groups. So far, primer cocktails have been rarely used, but might provide extremely comprehensive information on biodiversity. For instance, Govender, Singh, Groeneveld, Pillay, and Willows-Munro (2022) used six primer cocktails, each amplifying a different fragment of the COI-5P gene region, to explore the diversity of marine zooplankton. By combining primers optimized for different phyla, they were able to characterize at high resolution the diversity of the major taxonomic groups, including crustaceans, fish, echinoderms, mollusks, cnidarians and more. Govender et al. (2022) included up to four different reverse primers within the same PCR reaction, all targeting the same DNA fragment. However, in principle an even larger number of primers could be combined, to maximize the number of taxa that are amplified at high resolution, and the multiplex might include primers targeting different genomic regions, if they have comparable performance (see below). Such multiplexes including a large number of markers might boost the number of taxa amplified at high resolution, efficiently exploiting the available template DNA while limiting costs.

Nevertheless, this approach still needs major methodological developments. Primers often show strong variation in amplification efficiency, and the eDNA of different taxa normally is found at different concentrations. In standard PCRs, this is taken into account by tuning key parameters (e.g. number of cycles), but in a multiplex all the primers undergo the same number of cycles, therefore the mix should ideally include primers with comparable amplification performance, and targeting taxa with similar DNA concentration. Preliminary analyses can assess the similarity of primers, for instance checking via qPCR if they show analogous amplification patterns under the same conditions. Alternatively, multiplexes including markers with different efficiency and / or abundance of template DNA can be optimized by increasing the concentration of the primers with lower performance. Furthermore, designing a multiplex requires the identification of primers with similar annealing temperatures, but amplifying complementary groups. Specific bioinformatics tools have boosted our ability to identify the most appropriate metabarcoding primers (e.g. Riaz et al., 2011), but designing a multiplex will certainly need further developments for both bioinformatics and wet lab. Finally, current popular bioinformatics pipelines are optimized to process one marker at a time, and specific developments can be required to retrieve information from multiple metabarcodes from the same study (Porter & Hajibabaei, 2022).

2.5 Shotgun sequencing

Shotgun sequencing and other metagenomics approaches can extract a huge amount of information from the environmental DNA, and potentially allow the reconstruction of the whole community, without targeting a specific group (Gusareva et al., 2019; Parducci et al., 2019; Pedersen et al., 2016; Wang et al., 2021). In principle, this approach should bypass the DNA barcode amplification bias, might allow use the whole DNA available in the environment, providing information on all the trophic layers, and can even help to obtain information on the relative abundance of present taxa (Garrido-Sanz, Senar, & Piñol, 2022; Parducci et al.,

2017), thus overcoming many of the limitations associated to standard DNA metabarcoding.

Nevertheless, several issues continue to limit the broad-scale application of shotgun sequencing compared to the more standard metabarcoding. First, shotgun sequencing is much more expensive than PCR-based metabarcoding, and the associated bioinformatics pipelines remain complex. Furthermore, taxonomic identification relies on the existence of complete genomic databases. Unfortunately, so far genomic information outside the barcode regions is mostly limited to vertebrates, some plants (Alsos et al., 2020), and commercially important species. As a consequence, evidences of the advantage of shotgun sequencing over PCR-based metabarcoding for broad-scale environmental analyses remain mixed, so far (Bell et al., 2021; Parducci et al., 2019; Paula et al., 2022). Despite these issues, the continuing advances of sequencing and bioinformatics technologies suggest that shotgun will play an increasingly important role in the analysis of community-level variation, particularly for topical study systems such as ancient eDNA (Pedersen et al., 2016; Wang et al., 2021).

3. CONCLUSION: CHALLENGES AND OPPORTUNITIES FOR AN ALL-INCLUSIVE COMMUNITY ECOLOGY USING METABARCODING

One decade of advances on eDNA metabarcoding has fostered our ability to obtain biodiversity data, filling long-standing gaps on many components of both terrestrial and aquatic environments. Nevertheless, just a few studies have taken the challenge of attempting analyses covering multiple taxonomic groups, and trying to identify the complex multi-trophic interactions between them (but see Bloor et al., 2021; Calderón-Sanou et al., 2021; Calderón-Sanou et al., 2022; Martinez-Almoyna et al., 2019). Several approaches can now allow an all-inclusive community ecology, potentially allowing unprecedented understanding of patterns and processes underlying biodiversity variation, but both technical and conceptual developments will be required for a more widespread application of the all-inclusive ecology, and some challenges are shared by most approaches. So far, strong efforts have been devoted to the development of massive databases for standard barcodes, but just one or a few barcodes are unlikely to be enough to enable the characterization of the whole community. New reference databases can be generated using high-throughput sequencing approaches (e.g. genome skimming) that would allow covering broad sections of the genome (i.e. organelle(s) and nuclear ribosomal DNA), and might even serve as starting point for the identification of new markers (Coissac, Hollingsworth, Lavergne, & Taberlet, 2016). Furthermore, analyses of biotic interactions involving a large number of taxa remain extremely challenging. Novel frameworks have been proposed during the last years for the multi-trophic and multi-taxa analysis of communities, but a lot of work remains to be done to assess their power, strengths and limitations (e.g. Burian et al., 2021; D'Amen, Mod, Gotelli, & Guisan, 2018).

It is now clear that ongoing global changes determine very intricate effects on the organisms. For instance, species responses to climate change often alter the existing biotic interactions, and predicting a specie's response while ignoring interactions with its predators, foodsource or pathogen can lead to highly biased results (Sirén, Sutherland, Karmalkar, Duveneck, & Morelli, 2022; Urban et al., 2016). Yet, accounting for species interactions requires well-resolved information that is often missing, and that cover a large subset of existing biodiversity (Gilman, Urban, Tewksbury, Gilchrist, & Holt, 2010; Urban et al., 2016). DNA metabarcoding can heavily contribute to such endeavors of biodiversity studies, and we hope that methodological and conceptual advances, allowing an all-inclusive community ecology, will remain an active research area in the near future.

ACKNOWLEDGMENTS

GFF is funded by the European Research Council under the European Community's Horizon 2020 Programme, Grant Agreement no. 772284 (IceCommunities).

REFERENCES

Alsos, I. G., Lavergne, S., Merkel, M. K. F., Boleda, M., Lammers, Y., Alberti, A., . . . Coissac, E. (2020). The treasure vault can be opened: Large-scale genome skimming works well using herbarium and silica gel dried material. *Plants, 9* (4). doi:10.3390/plants9040432

Bell, K. L., Petit, R. A., Cutler, A., Dobbs, E. K., Macpherson, J. M., Read, T. D., . . . Brosi, B. J. (2021). Comparing whole-genome shotgun sequencing and DNA metabarcoding approaches for species identification and quantification of pollen species mixtures. *Ecology and Evolution, 11* (22), 16082-16098. doi:10.1002/ece3.8281

Bloor, J. M. G., Si-Moussi, S., Taberlet, P., Carrère, P., & Hedde, M. (2021). Analysis of complex trophic networks reveals the signature of land-use intensification on soil communities in agroecosystems.*Scientific Reports, 11* (1), 18260. doi:10.1038/s41598-021-97300-9

Burian, A., Mauvisseau, Q., Bulling, M., Domisch, S., Qian, S., & Sweet, M. (2021). Improving the reliability of eDNA data interpretation.*Molecular Ecology Resources, 21* (5), 1422-1433. doi:*https://doi.org/10.1111/1755-0998.13367*

Calderón-Sanou, I., Münkemüller, T., Zinger, L., Schimann, H., Yoccoz, N. G., Gielly, L., . . . Thuiller, W. (2021). Cascading effects of moth outbreaks on subarctic soil food webs. *Scientific Reports, 11* (1), 15054. doi:10.1038/s41598-021-94227-z

Calderón-Sanou, I., Zinger, L., Hedde, M., Martinez-Almoyna, C., Saillard, A., Renaud, J., . . . Thuiller, W. (2022). Energy and physiological tolerance explain multi-trophic soil diversity in temperate mountains. *Diversity and Distributions ,https://doi.org/10.1111/ddi.13529.* doi:*https://doi.org/10.1111/ddi.13529*

Coissac, E., Hollingsworth, P. M., Lavergne, S., & Taberlet, P. (2016). From barcodes to genomes: extending the concept of DNA barcoding.*Molecular Ecology, 25* , 1423-1428. doi:10.1111/mec.13549

D'Amen, M., Mod, H. K., Gotelli, N. J., & Guisan, A. (2018). Disentangling biotic interactions, environmental filters, and dispersal limitation as drivers of species co-occurrence. *Ecography, 41* , 1233-1244.

Ficetola, G. F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., . . . Taberlet, P. (2015). Replication levels, false presences, and the estimation of presence / absence from eDNA metabarcoding data. *Molecular Ecology Resources, 15* , 543–556. doi:DOI: 10.1111/1755-0998.12338

Floyd, R., Abebe, E., Papert, A., & Blaxter, M. (2002). Molecular barcodes for soil nematode identification. *Molecular Ecology, 11* , 839–850.

Garrido-Sanz, L., Senar, M. À., & Piñol, J. (2022). Relative species abundance estimation in artificial mixtures of insects using mito-metagenomics and a correction factor for the mitochondrial DNA copy number. *Molecular Ecology Resources, 22* (1), 153-167. doi:*https://doi.org/10.1111/1755-0998.13464*

Gilman, S. E., Urban, M. C., Tewksbury, J., Gilchrist, G. W., & Holt, R. D. (2010). A framework for community interactions under climate change. *Trends in Ecology & Evolution, 25* (6), 325-331. doi:10.1016/j.tree.2010.03.002

Govender, A., Singh, S., Groeneveld, J., Pillay, S., & Willows-Munro, S. (2022). Metabarcoding analysis of marine zooplankton confirms the ecological role of a sheltered bight along an exposed continental shelf.*Molecular Ecology, n/a* (n/a). doi:*https://doi.org/10.1111/mec.16567*

Guardiola, M., Uriz, M. J., Taberlet, P., Coissac, E., Wangensteen, O. S., & Turon, X. (2015). Deep-Sea, Deep-Sequencing: Metabarcoding Extracellular DNA from Sediments of Marine Canyons. *PLoS ONE, 10* (10), e0139633. doi:10.1371/journal.pone.0139633

Guerrieri, A., Carteron, A., Bonin, A., Marta, S., Ambrosini, R., Caccianiga, M., . . . Ficetola, G. F. (2022). Metabarcoding data reveal vertical multi-taxa variation in topsoil communities during the colonization of deglaciated forelands. *Molecular Ecology, in press* , DOI: 10.1111/mec.16669. doi:10.1111/mec.16669

Gusareva, E. S., Acerbi, E., Lau, K. J. X., Luhung, I., Premkrishnan, B. N. V., Kolundžija, S., . . . Schuster, S. C. (2019). Microbial communities in the tropical air ecosystem follow a precise diel cycle.*Proceedings of the National Academy of Sciences, 116* (46), 23299-23308. doi:10.1073/pnas.1908493116

Hebert, P. D. N., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B-Biological Sciences, 270* , 313-321.

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H., & Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator* . *Proceedings of the National Academy of Sciences of the United States of America, 101* , 14812-14817.

Hintikka, S., Carlsson, J. E. L., & Carlsson, J. (2022). The bacterial hitchhiker's guide to COI: Universal primer-based COI capture probes fail to exclude bacterial DNA, but 16S capture leaves metazoa behind.*Metabarcoding and Metagenomics, 6* , e80416.

Holman, L. E., de Bruyn, M., Creer, S., Carvalho, G., Robidart, J., & Rius, M. (2021). Animals, protists and bacteria share marine biogeographic patterns. *Nature Ecology & Evolution, 5* (6), 738-746. doi:10.1038/s41559-021-01439-7

Jamy, M., Biwer, C., Vaulot, D., Obiol, A., Jing, H., Peura, S., . . . Burki, F. (2022). Global patterns and rates of habitat transitions across the eukaryotic tree of life. *Nature Ecology & Evolution* . doi:10.1038/s41559-022-01838-4

Jurburg, S. D., Keil, P., Singh, B. K., & Chase, J. M. (2021). All together now: Limitations and recommendations for the simultaneous analysis of all eukaryotic soil sequences. *Molecular Ecology Resources, 21* (6), 1759-1771. doi:*https://doi.org/10.1111/1755-0998.13401*

Lawton, J. H., Bignell, D. E., Bolton, B., Bloemers, G. F., Eggleston, P., Hammond, P. M., . . . Laresen, T. B. (1998). Biodiversity inventories, indicator taxa and effects of habitat modification on tropical forest. *Nature, 391* , 72-76.

Li, F., Qin, S., Wang, Z., Zhang, Y., & Yang, Z. (2023). Environmental DNA metabarcoding reveals the impact of different land use on multitrophic biodiversity in riverine systems. *Science of the Total Environment, 855* , 158958. doi:*https://doi.org/10.1016/j.scitotenv.2022.158958*

Martinez-Almoyna, C., Thuiller, W., Chalmandrier, L., Ohlmann, M., Foulquier, A., Clément, J.-C., . . . Münkemüller, T. (2019). Multi-trophic β-diversity mediates the effect of environmental gradients on the turnover of multiple ecosystem functions. *Functional Ecology, 33* (10), 2053-2064. doi:*https://doi.org/10.1111/1365-2435.13393*

Parducci, L., Alsos, I. G., Unneberg, P., Pedersen, M. W., Han, L., Lammers, Y., . . . Wohlfarth, B. (2019). Shotgun Environmental DNA, Pollen, and Macrofossil Analysis of Lateglacial Lake Sediments From Southern Sweden. *Frontiers in Ecology and Evolution, 7* (189). doi:10.3389/fevo.2019.00189

Parducci, L., Bennet, K. D., Ficetola, G. F., Alsos, I. G., Suyama, Y., Wood, J. R., & Pedersen, M. W. (2017). Ancient plant DNA in lake sediments. *New Phytologist, 214* , 924-942.

Paula, D. P., Barros, S. K. A., Pitta, R. M., Barreto, M. R., Togawa, R. C., & Andow, D. A. (2022). Metabarcoding versus mapping unassembled shotgun reads for identification of prey consumed by arthropod epigeal predators. *GigaScience, 11* . doi:10.1093/gigascience/giac020

Pedersen, M. W., Ruter, A., Schweger, C., Friebe, H., Staff, R. A., Kjeldsen, K. K., . . . Willerslev, E. (2016). Postglacial viability and colonization in North America's ice-free corridor. *Nature, 537* , 45-49. doi:10.1038/nature19085

Porter, T. M., & Hajibabaei, M. (2022). MetaWorks: A flexible, scalable bioinformatic pipeline for high-throughput multi-marker biodiversity assessments. *PLoS ONE, 17* (9), e0274260. doi:10.1371/journal.pone.0274260

Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., & Coissac, E. (2011). ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research, 39* , e145. doi:10.1093/nar/gkr732

Rubinoff, D., Cameron, S., & Will, K. (2006). Are plant DNA barcodes a search for the Holy Grail? *Trends in Ecology & Evolution, 21* , 1-2.

Sirén, A. P. K., Sutherland, C. S., Karmalkar, A. V., Duveneck, M. J., & Morelli, T. L. (2022). Forecasting species distributions: Correlation does not equal causation. *Diversity and Distributions, 28* (4), 756-769. doi:*https://doi.org/10.1111/ddi.13480*

Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018).*Environmental DNA for biodiversity research and monitoring* . Oxford: Oxford University Press.

Urban, M. C., Bocedi, G., Hendry, A. P., Mihoub, J. B., Pe'er, G., Singer, A., . . . Travis, J. M. J. (2016). Improving the forecast for biodiversity under climate change. *Science, 353* (6304).

Wang, Y., Pedersen, M. W., Alsos, I. G., De Sanctis, B., Racimo, F., Prohaska, A., . . . Willerslev, E. (2021). Late Quaternary dynamics of Arctic biota from ancient environmental genomics. *Nature* . doi:10.1038/s41586-021-04016-x

DATA ACCESSIBILITY

All the relevant data are provided as supplementary material (Table S1).

AUTHOR CONTRIBUTIONS

The two authors jointly designed the study. GFF drafted the first draft of the manuscript, with sustantial contribution of PT.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section

ORCID

Gentile Francesco Ficetola: *https://orcid.org/0000-0003-3414-5155*

Pierre Taberlet:*https://orcid.org/**0000-0002-3554-5954***

Figure legends

Figure 1. Number and typologies of markers analyzed in 70 papers published in 2021-22 in seven representative scientific journals. We considered papers extracted from the Web of Science using the search term "DNA metabarcoding" and analyzing biodiversity variation. "Generalist markers" are markers that amplify multiple distantly related phyla and / or an entire domain of life, while studies focusing on "specific taxa" focus on a given taxonomic group (phylum, super-phylum or finer). Note that some studies focused on one specific taxon (e.g. fish), but used more than one marker to improve coverage. Additional details are provided in the Supplementary Methods and in Supplementary Table S1.

Table 1. Summary of approaches for all-inclusive community ecology, with examples of their strengths and limitations.

Approach	example	pros	cons
Combining many metabarcodes in the same study	Li et al. (2023) Analyzed freshwater biodiversity using four primers, focusing on bacteria; micro-eukaryotes; insects and fish	Good coverage of biodiversity Resolution can be high for the selected taxa	Costly Some taxon will always be missing

Universal markers	Holman et al. (2021) performed a joint biogeographical analysis of marine animals, protists and bacteria	Relatively cheap In principle, might cover the whole tree of life	Amplification rate and resolution are often heterogeneous across taxa
Combining universal and specific metabarcodes	Bloor et al. (2021) combined three universal (bacteria, eukaryotes, fungi) and four specific (seed plants, insects, springtails and earthworms) markers for a multi-trophic analysis of soil diversity	Good information on key groups Reduces the number of unrepresented taxa	Costly Resolution can be strongly heterogeneous across taxa
Multiplex of primers	Govender et al. (2022) used six primer cocktails to analyze the diversity of 14 zooplankton taxa	Potentially excellent resolution Potentially excellent coverage of the tree of life Cheaper than analyzing each taxon separately	Methodological developments required to optimize the multiplex Bioinformatics challenges
Shotgun sequencing	Pedersen et al. (2016) used ancient DNA to reconstruct post-glacial colonization patterns of plants, mammals and fish	Bypasses many limitations of metabarcoding (amplification, abundance) Can exploit the whole genomic DNA Can cover the whole tree of life Allows authentication of ancient eDNA	Assignation heavily depends on reference databases Very costly Complex analytical pipelines