MOLECULAR ECOLOGY

Molecular Ecology (2012) **21**, 1789–1793

INTRODUCTION

Environmental DNA

PIERRE TABERLET,* ERIC COISSAC,* MEHRDAD HAJIBABAEI† and LOREN H. RIESEBERG,‡§

**Laboratoire d'Ecologie Alpine, CNRS UMR 5553, Université Joseph Fourier, BP 53, F-38041 Grenoble Cedex 9, France, †Biodiversity Institute of Ontario, Department of Integrative Biology, University of Guelph, Guelph, Ontario, N1G 2W1, Canada, ‡Department of Botany, University of British Columbia, Vancouver, BC V6T 1Z4, Canada, §Department of Biology, Indiana University, Bloomington, IN 47405, USA*

Keywords: DNA metabarcoding, environmental DNA (eDNA), next-generation sequencing

Received 19 February 2012; revision accepted 19 February 2012

The emergence of environmental DNA in ecology

The first reference to environmental DNA (eDNA) dates back to 1987 and concerns a method for extracting microbial DNA from sediments (Ogram *et al.* 1987). However, the term really emerged at the beginning of the 2000s, mainly in the community of microbiologists (e.g. Rondon *et al.* 2000; Handelsman 2004). Environmental DNA refers to DNA that can be extracted from environmental samples (such as soil, water or air), without first isolating any target organisms. It is characterized by a complex mixture of genomic DNA from many different organisms and by possible degradation (i.e. DNA molecules are cut into small fragments). Total eDNA contains cellular DNA originating from living cells or organisms, and extracellular DNA (Levy-Booth *et al.* 2007; Pietramellara *et al.* 2009) resulting from natural cell death and subsequent destruction of cell structure.

Microbiologists have been analysing eDNA from soil or water for more than a decade, which has given them access to the genetics of uncultivable microorganisms. They had three main objectives: (i) identifying microbial taxa present in environmental samples; (ii) identifying the most important biochemical functions via the analysis of coding genes; and (iii) assembling whole genomes of uncultivable microorganisms. The analysis of eDNA by microbiologists was termed 'metagenomics' despite its departure from the initial definition of metagenomics (Handelsman *et al.* 1998), which only concerned assembly and functional analyses (objectives ii and iii) of eDNA from shotgun sequence data, but not identifications based on amplification and sequencing of marker genes such as 16S rDNA (objective i). The ambiguity arose from a seminal paper with 'metagenomics' in the title (Tringe *et al.* 2005) that combined shotgun sequencing with 16S identification.

DNA-based taxon identification has since been extended to the meiofauna (for example nematodes; Bhadury *et al.* 2006) and to macroorganisms, either using a PCR/cloning approach (Willerslev *et al.* 2003, 2007) or PCR/next-generation sequencing (Ficetola *et al.* 2008; Haile *et al.* 2009), but with the single goal of identifying the different organisms that contributed DNA to the sample. Clearly, the emergence of eDNA in broader ecological studies is linked to the availability of next-generation sequencers (NGS), which make it possible to bypass the expensive and time-consuming step of cloning and sequencing PCR products using Sanger sequencing. There is no doubt that the analysis of eDNA will soon be integrated in more and more ecological studies.

As new research fields emerge, different terminologies may be used by different research groups to describe somewhat similar approaches. A common vocabulary is necessary for unifying a nascent scientific community, at least by facilitating bibliographic surveys. For example, the term 'DNA barcoding' has been widely used for species-level DNA-based identification using standard markers. On the other hand, several different terms have been coined to describe the PCR-based approach for simultaneously identifying multiple taxa: ecometagenetics (Porazinska *et al.* 2010), ecogenomics (Chariton *et al.* 2010), environmental barcoding (Hajibabaei *et al.* 2011) or DNA metabarcoding (Pompanon *et al.* 2011; Riaz *et al.* 2011; plus 11 papers in this issue). DNA-based taxon identification methods mainly differ in three ways. First, the type of marker used is either one of the standardized DNA barcodes (a 658-bp region of the mitochondrial cytochrome c oxidase I gene for animals (Hebert *et al.* 2003), and two 500–800-bp plastid fragments of the large subunit of ribulose 1,5-bisphosphate carboxylase gene (*rbc*L) and the maturase K gene (*mat*K) for plants (CBoL Plant Working Group 2009)) or a much shorter DNA fragment, a mini-barcode (Hajibabaei *et al.* 2006), which permits the amplification of degraded DNA (e.g. Andersen *et al.* 2012; Bienert *et al.* 2012). Second, the identification level can vary among studies: either the marker used allows species-level identification, or it can only identify higher taxonomic levels such as genera, families or orders. Finally, the methodology employed must necessarily be adjusted based on the complexity and the degradation level of the DNA extract. Although it is possible to use standardized DNA barcodes for DNA extracts originating from single specimens or from bulk samples, only very short markers can be applied when dealing with degraded eDNA. Fig. 1 illustrates the possible types of

Correspondence: Pierre Taberlet;
E-mail: pierre.taberlet@ujf-grenoble.fr

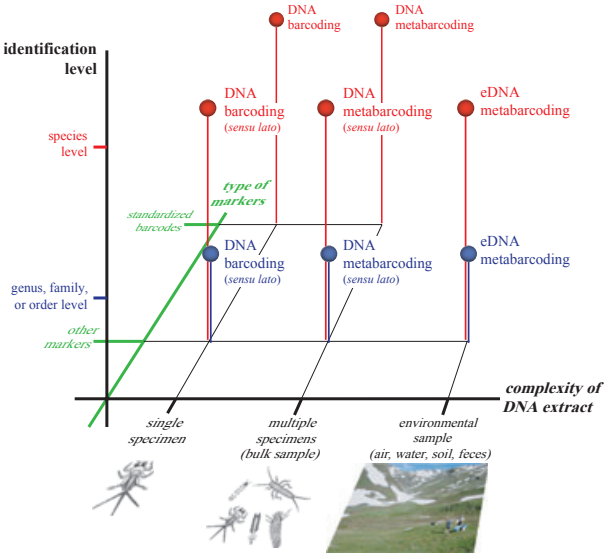

Fig. 1 Suggested terminology for DNA-based taxa identification according to the type of marker used, the level of identification and the complexity of template DNA.

DNA-based taxon identification and terminologies used according to these three factors. By analogy with well-established terms such as metagenomics, metatranscriptomics and metaproteomics, the term 'DNA metabarcoding' has been used in several manuscripts in this issue for all situations where multiple taxa are identified via a single experiment, or possibly 'eDNA metabarcoding' when using eDNA. However, because the term DNA barcoding aims for species-level DNA-based identification using standard markers, DNA metabarcoding as defined and used in this issue diverges from the standard DNA barcoding approach.

This issue

Methodological and technical advances

As mentioned earlier, technological advances, especially the introduction of NGS technologies, have been essential in the analysis of environmental DNA. Four articles in this issue are focused on different technological aspects of the analysis of environmental DNA. Shokralla *et al.* (2012) provide a thorough review of available NGS technologies and discuss different aspects of their chemistry, workflow and sequencing capabilities in the context of analysis of environmental DNA. Although NGS technologies are relatively new, there are already several platforms available commercially; understanding the pros and cons of each platform in the context of a particular analysis is key for their successful use. Additionally, Shokralla *et al.* (2012) provide an overview of different modifications and additional tools such as oligonucleotide tags (for combining several environmental samples in one sequence run) or capture arrays (to circumvent amplification biases).

Selective amplification and sequencing of rare species has been challenging in environmental DNA analysis. Boessenkool *et al.* (2012) demonstrate a technical advance in the analysis of rare DNA templates by blocking abundant human contaminant DNA at the PCR step. This approach can facilitate the analysis of ancient DNA or in other cases where rare taxa in a mixture are the target of investigation.

Robust DNA extraction is a key step in the analysis of environmental DNA. Taberlet *et al.* (2012a) discuss a new protocol for preparing and extracting DNA from large amount of soil, targeting extracellular DNA. This method is based on using a saturated phosphate buffer, with the goal of providing a DNA extract that is as representative as possible of the local biodiversity. This protocol has the potential of allowing the analysis of many different groups of organisms using the same DNA extraction.

The implementation of large-scale eDNA-based ecological studies is highly dependent on the availability of suitable short metabarcode. Epp *et al.* (2012) developed such metabarcodes for fungi, bryophytes, enchytraeids, beetles and birds. These new markers have been evaluated both *in silico* and *in vitro*, focusing not only on taxonomic resolution but also on the effectiveness of the primers in minimizing amplification biases among different target taxa.

Bioinformatic aspects

Taxonomic identification using DNA relies on a variety of sequence analysis techniques. When this approach is coupled with NGS for analysing biodiversity from eDNA, bioinformatics becomes even more important. Until very recently, most of the work published on eDNA concerns microorganisms. Thus, most of the bioinformatics tools developed are devoted to the analysis of microbial communities. In their review, Coissac *et al.* (2012) reviewed the set of bioinformatic tools published by the DNA barcoding community, by environmental microbiologists and by scientists working on plant and animal to present a subset of tools and methods applicable for high-throughput plant and animal DNA-based identification. Each step of a DNA metabarcoding experiment is considered, from the primer evaluation and design, to the data analysis.

Nonetheless, one of the main bioinformatic challenges that remain is the transformation of a list of sequences into a list of species. Two papers of this special issue deal with this problem from two different points of view. With the Automatic Barcode Gap Discovery method, Puillandre *et al.* (2012) propose a new approach for delimiting species from a set of DNA barcode sequences. Even if this method can be considered as mainly devoted to the taxonomical side of the DNA barcoding, it can be also considered as an interesting approach for building molecular operational taxonomic units when no reference database is available for analysing eDNA metabarcoding sequences. Conversely, the method proposed by Zhang *et al.* (2012) requires a reference database, but by relying on the fuzzy set theory, a complete reference database is unnecessary. This is a useful property because a large and unknown portion of extant biodiversity remains uncharacterized.

Microbial ecology

Because microbial diversity is so high and many bacterial and fungal species cannot be cultivated, environmental microbiologists were the first to develop an eDNA metabarcoding approach. This scientific community can be approximately divided between scientists working on soil and those working on aquatic environments. Zinger *et al.* (2012) focused their review on aquatic ecosystems. Considering the switch from Sanger sequencing to NGS technologies, this review presents advances, gaps and caveats associated with these approaches. This included an evaluation of the power and limitations of the available methodologies, from water sampling to sequence analysis.

Three other articles also deal with microorganisms. In the first one, Blaalid *et al.* (2012) present results on the ectomycorrhizal fungal communities associated with *Bistorta vivipara* along a primary succession gradient, showing that fungal diversity increases significantly towards the climax vegetation. As the description of microorganism communities becomes increasingly easy to obtain, the next research step will be to link them with the functional characteristics of the ecosystem. Barberán *et al.* (2012) is one of the first papers to link functional traits and DNA metabarcoding results. Their analysis of a subset of the Global Ocean Sampling project (Rusch *et al.* 2007) demonstrates that some simple characteristics of the sequences can be used as functional traits, making it feasible to categorize samples without a full taxonomic inventory.

Stoof-Leichsenring *et al.* (2012) analysed diatoms from core sediments of tropical lakes in Kenya. They first developed a new approach based on the amplification and sequencing of short *rbc*L fragments to identify the different diatoms. This DNA-based identification method showed a significant correlation with morphological identifications and revealed cryptic lineages that were undetected by morphology. Temporal variations in diatom assemblages during the last two centuries were precisely analysed, showing the potential of this DNA-based approach for assessing past environmental factors.

Diet analysis

The availability of next-generation sequencing platforms has also boosted studies on diet analysis. DNA extracts from gut content or from faeces show many analogies with eDNA: they contain a mixture of genomic DNA from different organisms, and at least in faeces, DNA is highly degraded. As a consequence, at the technical level, analysing DNA from faeces requires the same approach as when analysing eDNA. Pompanon *et al.* (2012) reviewed the power and potential pitfalls of the DNA metabarcoding approach for analysing diets, giving to the readers all the necessary background for initiating such studies.

A second paper in this section deals with carnivore diets (Shehzad *et al.* 2012). It presents a simple methodology based on the use of vertebrate specific primers, a blocking oligonucleotide and NGS to analyse the eclectic vertebrate diet of the leopard cat (*Prionailurus bengalensis*), which includes fishes, amphibians, birds and mammals. With only a few adjustments of the blocking oligonucleotide, this approach has the potential of being widely used for studying the diet of carnivores that eat vertebrates.

Biodiversity assessment of plants and animals

Significant improvements have recently been made in this research area. By sampling eDNA from safari parks, zoological gardens and farms, Andersen *et al.* (2012) demonstrated that analyses of soil DNA can be used to record the presence of large mammals. They also provide useful information about the different factors influencing the success of the approach. Similarly, Bienert *et al.* (2012) showed that traditional and time-consuming earthworm inventories can be complemented by eDNA analysis, allowing large-scale studies of earthworm diversity. Surprisingly, the method developed in the French Alps has the potential to work on earthworm communities anywhere in the world.

The two next papers of this section deal with ancient DNA of plants isolated from permafrost samples. The DNA fragments isolated from frozen sediments can be considered fossils that provide information about past ecosystems. Jørgensen *et al.* (2012a) were able to reconstruct the past vegetation of a Greenland nunatak for the last 5000 years, suggesting that flora from small and isolated ice-free areas can react quickly to climate change. The goal of the second paper (Jørgensen *et al.* 2012b) was to compare eDNA analysis with more traditional methods (pollen and macrofossil analysis). The results showed that pollen, macrofossils and eDNA are complementary and that together they lead to a more precise reconstruction of plant palaeocommunities, confirming that large areas of Northern Hemisphere were ecologically stable during the Late Pleistocene.

The paper of Hiiesalu *et al.* (2012) is not based on the analysis of eDNA, but rather on the analysis of bulk root samples. Both types of approaches require the simultaneous identification of many taxa with a single DNA amplification, but the main difference is that bulk samples contain less degraded DNA than eDNA. The objective of the authors was to compare estimations of plant species richness using classical above-ground surveys with DNA-based below-ground surveys. It appeared that the diversity revealed by DNA-based below-ground surveys was higher than that from above-ground surveys.

Perspectives

Three perspective articles conclude the collection of articles in this special issue. Yoccoz (2012) provides a broad overview of the approaches used for environmental DNA analysis and their potential for future ecological research. The article discusses examples from different applications of environmental DNA analysis from diet analysis to understanding food web interactions, species distributions and

niches and biodiversity surveys. By discussing methodological/computational as well as societal challenges, Yoccoz (2012) provides a thorough review of pros and cons of using environmental DNA in ecology and concludes by cautioning the community in addressing a fundamental question: 'How do we integrate eDNA with more classical approaches so as to get more from both?'

The second perspective article by Baird & Hajibabaei (2012) introduces Biomonitoring 2.0, the use of environmental DNA information for biomonitoring and ecosystem assessment. The article explains approaches used today by environmental agencies for ecosystem assessment. By providing limitations of current approaches, Baird & Hajibabaei (2012) set the stage for identifying bioindicator taxa such as aquatic macroinvertebrates (currently used in biomonitoring) directly from the analysis of DNA barcodes from environmental DNA. Additionally, the Biomonitoring 2.0 scheme should lead to environmental assessment using a much wider array of species besides typical bioindicators.

The third perspective article by Taberlet *et al.* (2012b) introduces DNA metabarcoding as an umbrella approach for the use of environmental DNA in analysis of taxa in ecological research. The article discusses different aspects of DNA-based species identification work and argues that standard DNA barcoding based on the analysis of single specimens has limitations in high-throughput analysis and for applications involving degraded and fragmented DNA in mixed environmental samples. The use of NGS tools with a wider array of DNA markers that can target degraded DNA is key to the application of a DNA metabarcoding approach in ecological investigations. Although the use of this more liberal criterion differentiates standard DNA barcoding from DNA metabarcoding, future technological advances should lead to better coverage of sequence information and less biased data for optimal applicability. As shown in several articles in this special issue, this approach opens up many possibilities today for the applicability of DNA information in ecology.

Acknowledgements

We warmly thank Hilary McNaughton, Adrienne Nye, Gregor Hutton and Tim Vines for their patience and efficient support during the editing and publishing process of this Special Issue, and all authors and reviewers for their valuable contributions and their efforts to meet deadlines.

References

Andersen K, Bird KL, Rasmussen M *et al.* (2012) Meta-barcoding of ''dirt'' DNA from soil reflects vertebrate biodiversity. *Molecular Ecology*, **21**, 1966–1979.

Baird D, Hajibabaei M (2012) Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Molecular Ecology*, **21**, 2039–2044.

Barberán A, Fernández-Guerra A, Bohannan BJM, Casamayor EO (2012) Exploration of community traits as ecological markers in microbial metagenomes. *Molecular Ecology*, **21**, 1909–1917.

Bhadury P, Austen MC, Bilton DT *et al.* (2006) Molecular detection of marine nematodes from environmental samples: overcoming eukaryotic interference. *Aquatic Microbial Ecology*, **44**, 97–103.

Bienert R, de Danieli S, Miquel C *et al.* (2012) Tracking earthworm communities from soil DNA. *Molecular Ecology*, **21**, 2017–2030.

Blaalid R, Carlsen T, Kumar S *et al.* (2012) Changes in the root-associated fungal communities along a primary succession gradient analysed by 454 pyrosequencing. *Molecular Ecology*, **21**, 1897–1908.

Boessenkool S, Epp L, Haile J *et al.* (2012) Blocking human contaminant DNA during PCR allows amplification of rare mammal species from sedimentary ancient DNA. *Molecular Ecology*, **21**, 1806–1815.

CBoL Plant Working Group (2009) A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 12794–12797.

Chariton AA, Court LN, Hartley DM, Colloff MJ, Hardy CM (2010) Ecological assessment of estuarine sediments by pyrosequencing eukaryotic ribosomal DNA. *Frontiers in Ecology and the Environment*, **8**, 233–238.

Coissac E, Riaz T, Puillandre N (2012) Bioinformatic challenges for DNA metabarcoding of plants and animals. *Molecular Ecology*, **21**, 1834–1847.

Epp LS, Boessenkool S, Bellemain EP *et al.* (2012) New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems. *Molecular Ecology*, **21**, 1821–1833.

Ficetola GF, Miaud C, Pompanon F, Taberlet P (2008) Species detection using environmental DNA from water samples. *Biology Letters*, **4**, 423–425.

Haile J, Froese DG, MacPhee RDE *et al.* (2009) Ancient DNA reveals late survival of mammoth and horse in interior Alaska. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 22352–22357.

Hajibabaei M, Smith MA, Janzen DH *et al.* (2006) A minimalist barcode can identify a specimen whose DNA is degraded. *Molecular Ecology Notes*, **6**, 959–964.

Hajibabaei M, Shokralla S, Zhou X, Singer GAC, Baird DJ (2011) Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS ONE*, **6**, e17497.

Handelsman J (2004) Metagenomics: application of genomics to uncultured microorganisms. *Microbiology and Molecular Biology Reviews*, **68**, 669.

Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chemistry and Biology*, **5**, R245–R249.

Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **270**, 313–321.

Hiiesalu I, Öpik M, Metsis M *et al.* (2012) Plant species richness belowground: higher richness and new patterns revealed by next-generation sequencing. *Molecular Ecology*, **21**, 2004–2016.

Jørgensen T, Kjær KH, Haile J *et al.* (2012a) Islands in the ice: detecting past vegetation on Greenlandic nunataks using historical records and sedimentary ancient DNA meta-barcoding. *Molecular Ecology*, **21**, 1980–1988.

Jørgensen T, Haile J, Möller P *et al.* (2012b) A comparative study of ancient sedimentary DNA, pollen and macrofossils from permafrost sediments of northern Siberia reveals long-term vegetational stability. *Molecular Ecology*, **21**, 1989–2003.

Levy-Booth DJ, Campbell RG, Gulden RH *et al.* (2007) Cycling of extracellular DNA in the soil environment. *Soil Biology and Biochemistry*, **39**, 2977–2991.

Ogram A, Sayler GS, Barkay T (1987) The extraction and purification of microbial DNA from sediments. *Journal of Microbiological Methods*, **7**, 57–66.

Pietramellara G, Ascher J, Borgogni F *et al.* (2009) Extracellular DNA in soil and sediment: fate and ecological relevance. *Biology and Fertility of Soils*, **45**, 219–235.

Pompanon F, Coissac E, Taberlet P (2011) Metabarcoding a new way to analyze biodiversity. *Biofutur*, **3**, 30–32.

Pompanon F, Deagle BE, Symondson WOC *et al.* (2012) Who is eating what: diet assessment using next generation sequencing. *Molecular Ecology*, **21**, 1931–1950.

Porazinska DL, Giblin-Davis RM, Esquivel A *et al.* (2010) Ecometagenetics confirms high tropical rainforest nematode diversity. *Molecular Ecology*, **19**, 5521–5530.

Puillandre N, Lambert A, Brouillet S, Achaz G (2012) ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, **21**, 1864–1877.

Riaz T, Shehzad W, Viari A *et al.* (2011) ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, **39**, e145.

Rondon MR, August PR, Bettermann AD *et al.* (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Applied and Environmental Microbiology*, **66**, 2541–2547.

Rusch DB, Halpern AL, Sutton G *et al.* (2007) The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biology*, **5**, e77.

Shehzad W, Riaz T, Nawaz MA *et al.* (2012) Carnivore diet analysis based on next generation sequencing: application to the leopard cat (*Prionailurus bengalensis*) in Pakistan. *Molecular Ecology*, **21**, 1951–1965.

Shokralla S, Spall J, Gibson J, Hajibabaei M (2012) Next-generation sequencing technologies for environmental DNA research. *Molecular Ecology*, **21**, 1794–1805.

Stoof-Leichsenring KR, Epp LS, Trauth MH, Tiedemann R (2012) Hidden diversity in diatoms of Kenyan Lake Naivasha: a genetic approach detects temporal variation. *Molecular Ecology*, **21**, 1918–1930.

Taberlet P, Prud'homme S, Campione E *et al.* (2012a) Soil sampling and isolation of extracellular DNA from large amount of starting material suitable for metabarcoding studies. *Molecular Ecology*, **21**, 1816–1820.

Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012b) Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, **21**, 2045–2050.

Tringe SG, von Mering C, Kobayashi A *et al.* (2005) Comparative metagenomics of microbial communities. *Science*, **308**, 554–557.

Willerslev E, Hansen AJ, Binladen J *et al.* (2003) Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science*, **300**, 791–795.

Willerslev E, Cappellini E, Boomsma W *et al.* (2007) Ancient biomolecules from deep ice cores reveal a forested Southern Greenland. *Science*, **317**, 111–114.

Yoccoz NG (2012) The future of environmental DNA in ecology. *Molecular Ecology*, **21**, 2031–2038.

Zhang A-B, Muster C, Liang H-B *et al.* (2012) A fuzzy-set-theory-based approach to analyse species membership in DNA barcoding. *Molecular Ecology*, **21**, 1848–1863.

Zinger L, Gobet A, Pommiers T (2012) Two decades of describing the unseen majority of aquatic microbial diversity. *Molecular Ecology*, **21**, 1878–1896.

doi: 10.1111/j.1365-294X.2012.05542.x