Environmental (e)DNA: what’s behind the term? Clarifying the terminology and recommendations for its future use in biomonitoring

An opinion paper

Jan Pawlowski 1,2,3, Laure Apothéloz-Perret-Gentil 1,2, Florian Altermatt 4,5

1 Department of Genetics and Evolution, University of Geneva, 1211 Geneva, Switzerland
2 ID-Gene ecodiagnostics, Campus Biotech Innovation Park, 1202 Geneva, Switzerland
3 Institute of Oceanology, Polish Academy of Sciences, 81-712 Sopot, Poland
4 Department of Aquatic Ecology, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland
5 Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zürich, Switzerland

This document is the accepted manuscript version of the following article:
Abstract

The last decade brought a spectacular development of so-called environmental (e)DNA studies. In general, “environmental DNA” is defined as DNA isolated from environmental samples, in opposition to genomic DNA that is extracted directly from specimens. However, the variety of different sources of eDNA and the range of taxonomic groups that are targeted by eDNA studies is large, which has led to some discussion about the breadth of the eDNA concept. In particular, there is a recent trend to restrict the use of the term “eDNA” to the DNA of macroorganisms, which are not physically present in environmental samples. In this paper, we argue that such a distinction may not be ideal, because the eDNA signal can come from organisms across the whole tree of life. Consequently, we advocate that the term “eDNA” should be used in its generic sense, as originally defined, encompassing the DNA of all organisms present in environmental samples, including microbial, meiofaunal, and macrobial taxa. We first suggest specifying the environmental origin of the DNA sample, such as water eDNA, sediment eDNA, or soil eDNA. A second specification would then define the taxonomic group targeted through PCR amplification, such as fish eDNA, invertebrate eDNA, bacterial eDNA. This terminology does also not require assumptions about the specific state of the DNA sampled (intracellular or extracellular). We hope that such terminology will help better define the scope of eDNA studies, especially for environmental managers, who use them as reference in routine biomonitoring and bioassessment.

The importance of environmental DNA studies

During the last decade we observe a rapidly increasing number of studies that are using DNA isolated from the environment, especially for aquatic ecosystems, both freshwater and marine. These ecosystems are under immense anthropogenic pressures, and the biodiversity and
associated ecosystem processes and services are heavily and negatively affected (Dudgeon, 2019; Reid et al., 2019). Consequently, effective management is needed, and this itself depends on accurate, timely, and reliable assessments of the state and change of the organismal communities, either by describing their biodiversity or using them for calculating indices as proxies describing the environmental state (Jackson et al., 2016; Pawlowski et al., 2018). A major limitation of past assessment methods, however, was their high cost, methodological diversity across taxonomic groups, as well as the inability to upscale the methods in time and space. Highly resolved biomonitoring data, however, are crucially needed, possibly depending on novel technologies. An example of such an advancement is the use of molecular techniques, and the study of eDNA in particular, which have been proposed to be a game-changer for bioassessment and monitoring of biodiversity (Altermatt et al., 2020; Deiner et al., 2017; Kelly et al., 2014; Pawlowski et al., 2018; Pfrender et al., 2010; Taberlet et al., 2018). Within few years, many studies on bioassessment in aquatic systems started to use and develop eDNA tools. However, the objectives, the methodologies, the source of eDNA, and the organisms targeted by these studies can be very different. In parallel, these novel techniques and the use of eDNA for bioassessment have raised high expectations, especially from stakeholders, and they are progressively implemented in ongoing biodiversity monitoring programs and bioassessment studies (e.g., Herder et al., 2014; Pawlowski et al., 2020; Thomsen & Willerslev, 2015). While many of the expectations are probably realistic, there is also regular misunderstandings and misconceptions on the potential but also limitations (or at least boundary-conditions) of eDNA studies, which is further fuelled by different uses of terminology in the research field itself.

Here, we advocate for a common terminology, which specifies where the DNA comes from (i.e., from which environment), and which organisms are looked for (i.e., which organisms’ DNA is targeted with PCR). Our proposed terminology links to the original definition of
“environmental DNA” (Taberlet et al., 2012). It would clarify issues concerning the state of the DNA sampled and would also resolve discussions about the inclusion/exclusion of certain organismal groups based on their size only.

The evolution of the eDNA concept: from microbial to macrobial studies

At the beginning, the technical concept of environmental DNA (eDNA) was used principally to explore microbial diversity. At its basis laid a ground-breaking idea that the analysis of RNA or DNA extracted from environmental samples can be used to assess the natural diversity of microorganisms (Pace et al., 1986). The early studies were based on RNA isolated from environmental samples (Olsen et al., 1986; Stahl et al., 1984). However, very rapidly environmental DNA became the focus of studies on microbial diversity and several papers were published on how to recover DNA from environmental samples (Ogram et al., 1987; Paul & Myers, 1982; Somerville et al., 1989; Steffan et al., 1988). At that time, the authors either used a descriptive term “the DNA isolated from environmental samples” (Sommerville et al., 1989) or specifically referred to targeted organisms quoting “bacterial DNA” (Steffan et al., 1988) or “microbial DNA” (Ogram et al., 1987; Paul & Myers, 1982). To our knowledge the term “environmental DNA” was used for the first time by Ogram et al. (1987) in a figure describing the protocol for the isolation of microbial DNA from sediments. Later, Somerville et al. (1989) used it in reference to the work of Pace et al. (1986).

The invention of PCR amplification (Saiki et al., 1988) contributed to the rapid development of studies exploring microbial diversity in environmental samples. These studies totally changed our perception of bacterial diversity, revealing huge numbers of uncultivable species in the ocean (Giovannoni et al., 1990; Schmidt et al., 1991) and in the soil (Picard et al. 1992; Torsvik et al., 1990). Yet, the term “environmental DNA” was only sporadically used in these
early days of bacterial diversity exploration (Suzuki et al., 1997). In parallel, the term was sometimes used to refer to the “free” DNA released by the lysis of dying cells (Kloos et al., 1994), corresponding to the “extracellular DNA”. However, the research focus of this and previous similar studies (Deflaun et al., 1986, Paul et al. 1987) was on the dynamics and biological potential of “extracellular DNA” rather than its use for biodiversity monitoring or bioassessment.

Since 2000, the term “environmental DNA” became much more commonly used (Fig. 1) in a variety of studies and often in a biodiversity context, such as in the description of new environmental microbial phyla (Huber et al., 2002), or the exploration of microbial diversity in the extreme environments (Gordon et al., 2000). The term was also generally adopted in early studies exploring microbial eukaryotes diversity (Bass & Cavalier-Smith, 2004; Berney et al., 2004; Holzmann et al., 2003). In all of these studies, the term did not make a specific assumption about the state of the DNA sampled, and was mostly linked to the microbial organisms being directly sampled. Shortly after, the scope of environmental DNA studies was expanded to the detection of large animals, such as fish or amphibians, whose DNA traces are preserved in water for a certain time, as demonstrated in a seminal paper by Ficetola et al. (2008). This new application constituted a turning point in the research field on environmental DNA, prompting a series of studies using eDNA to monitor biological invasions and/or endangered species in aquatic environments (e.g., Darling & Mahon 2011; Macher et al., 2018; Mächler et al., 2014; Thomsen et al., 2012). In parallel, the development of high-throughput sequencing opened new perspectives to apply the eDNA approach to survey the community of species at very high yield and relatively low cost compared to the traditional cloning approach.

Rapid increases of the number of eDNA studies in the first and second decades of the 21st Century called for clarification of eDNA terminology. This was done by Taberlet et al. (2012) in a special issue of Molecular Ecology, where eDNA was defined as “DNA that can be
extracted from environmental samples (such as soil, water or air), without first isolating any target organisms”. Importantly, this definition does not make assumptions on the state of the DNA sampled (extracellular or intracellular, tissue fragments, gametes, etc.), and is also not restricted to any specific group of organisms. The authors also clarified the differences between DNA barcoding, DNA metabarcoding (referring to analysis of bulk samples), and eDNA metabarcoding (defined as a study that allow identification of multiple taxa using environmental DNA as a template material) (Taberlet et al., 2012). It was also proposed to expand the concept of eDNA and to include DNA extracts from gut content or from faeces, both containing a mixture of genomic DNA from different organisms (Yoccoz et al., 2012). However, this latter suggestion was rarely followed by authors using DNA metabarcoding for diet analysis (Srivathsan et al., 2015).

The recent massive increase of eDNA studies applied to conservation biology and biodiversity assessments targeting mainly macro-organisms prompted some authors to redefine environmental DNA as “a mixture of potentially degraded DNA from many different organisms” (Bohmann et al., 2014; Cristescu, 2014) or “genetic material obtained directly from environmental samples (...) without any obvious signs of biological source material” (Thomsen & Willerslev, 2015). The latter definition was introduced to highlight the fact that the nature of DNA present as traces in environmental samples is different from the DNA derived from living microorganisms or meiofauna that can be present in eDNA samples (Goldberg et al., 2015). Although the authors assume that the macrobial eDNA exists predominantly inside mitochondria and cells (Turner et al., 2014), they state that part of it may have extracellular origin. The importance of subcellular origin of eDNA has also been suggested by other studies (Moushomi et al. 2019). This creates an additional confusion between the terms “environmental DNA” and “extracellular DNA”. The latter is also used for biodiversity surveys (Corinaldesi et al., 2018; Guardiola et al., 2015; Pearman et al., 2016;
Taberlet et al., 2012), and can be abbreviated as “eDNA”, although mainly in studies related to microbial biofilm formation (Harmsen et al., 2010).

Currently, two definitions of environmental DNA are used in ecological studies in parallel. On the one hand, the definition of eDNA *sensu lato* is used in global biodiversity surveys that analyse microbial, meiofauna and macrofauna communities, focusing on their ecological interactions (Deiner et al., 2016; Djurhuus et al., 2020; Zhang et al., 2020) and temporal and spatial dynamics (Altermatt et al., 2020; Bálint et al., 2018; Carraro et al., 2020). Such definition is also commonly used in environmental biomonitoring studies that target different groups of bioindicators to infer or predict biotic indices (Cordier et al., 2018, 2019; Li et al., 2018; Pawlowski et al., 2018; Stoeck et al., 2018). This definition also recognizes that samples of environmental DNA contain both intra- and extracellular DNA of microbial and macrobial species, and that the type of DNA captured may depend on capture method (Deiner et al., 2015). On the other hand, the definition of eDNA *sensu stricto* only referring to (mostly or even exclusively extracellular) DNA of macrobial organisms is especially used in conservation biology to monitor invasive and/or endangered species (Borrell et al., 2018; Lacoursière-Roussel & Deiner, 2019; O’Sullivan & al., 2020), as well as in ecology to survey animal and plants communities and to study biodiversity patterns in aquatic ecosystems (see Deiner et al., 2016; Nguyen et al., 2020). Sometimes, the eDNA concept also includes the DNA extracted from bulk samples (see e.g., Lynggaard et al. 2019, Nielsen et al. 2019, both working on bulk samples but published in a journal dedicated to environmental DNA). Hence, the same term is used in slightly different ways for different types of studies, which can lead to misunderstandings or confusion.

**Recommended eDNA terminology**
Here, we suggest maintaining the original concept of environmental DNA, defined as a total pool of DNA isolated from environmental samples (Taberlet et al., 2012, 2018). This general concept assumes that the environmental DNA is defined primarily by its origin and not by its taxonomic composition nor its specific structural state (intra- or extracellular). Indeed, such definition covers the DNA of various taxonomic origins, including living microorganisms and meiofauna-size taxa, as well as macrofauna traces, possible larval stages or gametes, as well as extracellular DNA. The eDNA can be isolated from various types of material, principally soil, sediment, and water, but also from air, biofilm and organic remains, such as faeces that may contain DNA of different origin. In principle, such definition precludes any type of size-based physical preselection of target taxa, such as sieving or kicknet sampling.

We think that restricting eDNA definition to the traces of large-size macro-organisms, which are not physically present in environmental DNA samples is unnecessary and possibly confusing. First and most importantly from a semantic point of view, such definition refers to the target DNA that is amplified from the environmental sample, not to the DNA that is isolated from the environment. It is incorrect to say that “macrobial DNA is isolated directly from an environmental sample”, as the separation between microbial and macrobial DNA occurs only later in processing of eDNA samples. It is often forgotten that the macrobial DNA represents only a very small fraction of the total DNA recovered from the environment, which is of microbial origins mainly (Stat et al. 2017). Secondly, such a definition does also not take in consideration the eggs, larvae or other small stages of macrofauna life cycles can be present in environmental samples and have been suspected to contribute to some of the eDNA signals observed. Finally, it implicitly assumes that the structural state of macrobial eDNA is different because it originates from DNA traces, while microbial or meiofaunal eDNA might derive from whole organisms. However, as demonstrated by numerous extracellular DNA studies, the microorganisms and meiofauna are represented there as much as macro-organisms.
We recommend that the eDNA studies adopt a **two level terminology** that clearly specifies the origin of environmental samples and the target taxa. At a first level, specific terms, such as water eDNA, sediment eDNA, or soil eDNA, would provide information about the environmental source of eDNA and/or the sampling methodology. Given that it is an environmental sample already, the “e” of the environment is basically defined by the specific environment given, and its use as a prefix of DNA might be considered as redundant but may be clearer. A second level of specification would then provide information about the taxonomic groups targeted in environmental samples, that is their DNA amplified and sequenced, such as fish eDNA, invertebrate eDNA, diatom eDNA, bacterial eDNA. In this case, the use of “eDNA” would clearly separate it from DNA extracted from tissues or cultures of these organisms, as well as from the DNA extracted from bulk samples that are not considered as eDNA here. Our proposed terminology does not require assumptions about the specific state of the DNA sampled (e.g., from cells, tissue fragments, gametes or free-floating), which is also generally not assessed, nor can it be told apart based on the sequence information. Obviously, this second level of the terminology is not required in the case of studies based on PCR-free approaches.

In our opinion, our proposed two-level terminology will contribute to clarify the scope of eDNA studies. Given the rapidly increasing number of studies using eDNA for biomonitoring and bioassessment, it is important to be as precise as possible regarding their objectives and outcomes, for example by specifying that the particular study was conducted using water eDNA and focusing on fish eDNA. A terminological clarity is particularly important for environmental managers who are not always aware about the various opportunities offered by new technology. Restricting the use of the term solely to tracing the large-sized organisms is drawing attention away from what we see as the most prominent application of eDNA technology, namely being a unique tool capable of providing a global assessment of ecological
status including different biological quality elements at a time. We are convinced that retaining
the original broad definition of eDNA highlighting its universal character will contribute to
expanding the field eDNA research and its successful application.

Acknowledgements

We thank Simon Creer and an anonymous reviewer for constructive comments on the
manuscript. Funding is from the Swiss National Science Foundation Grants No
31003A_173074 (FA) and No 31003A_179125 (JP), as well as from the University of Zurich
Research Priority Programme “URPP Global Change and Biodiversity” (to F.A.).
Fig. 1. The number of publications by years referring to environmental DNA studies targeting microbial diversity, macrobial diversity, or both. Microbial diversity encompasses bacterial and viral diversity as well as eukaryotic micro- and meiofauna. The figure is based on a PubMed NCBI search (on May 5th, 2020) of titles and abstract containing the term “Environmental DNA”, excluding studies containing “medical” or “cancer”. This resulted in 1009 papers. After manual inspection, 192 papers were removed from this list because they clearly were outside a biodiversity context. The full list of all papers considered is available upon request.
Fig. 2. Principal types of environmental samples and target taxonomic groups commonly used in biomonitoring and bioassessment. This figure shows the two levels at which environmental DNA term can be specified. The first level indicates where the DNA comes from (the type of environmental substrate sampled, such as soil, sediment, biofilm, or water). The second level is then specifying what taxonomic group is targeted by PCR amplification (based on the specific choice of primers), including bacteria, protists, fungi, diatoms, meiofauna, arthropods, molluscs, amphibians and fish. The width of the line corresponds qualitatively to the common usage of particular eDNA extracted for each taxonomic group. Other types of environments, such as air or faeces, were not included for simplicity.
References


Moushomi, R., Wilgar, G., Carvalho, G., Creer, S. & Seymour, M. (2019) Environmental DNA size sorting and degradation experiment indicates the state of Daphnia magna mitochondrial and nuclear eDNA is subcellular. Scientific Reports. 9, 12500.


