Ecosystems monitoring powered by environmental genomics: a review of current strategies with an implementation roadmap

Tristan Cordier, Laura Alonso-sáez, Laure Apothéloz-perret-gentil, Eva Aylagas, David Bohan, Agnes Bouchez, Anthony Chariton, Simon Creer, Larissa Frühe, François Keck, et al.

To cite this version:
Tristan Cordier, Laura Alonso-sáez, Laure Apothéloz-perret-gentil, Eva Aylagas, David Bohan, et al.. Ecosystems monitoring powered by environmental genomics: a review of current strategies with an implementation roadmap. Molecular Ecology, 2021, 30 (13), pp.2937-2958. 10.1111/mec.15472. hal-02860583

HAL Id: hal-02860583
https://hal.inrae.fr/hal-02860583
Submitted on 25 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Ecosystems monitoring powered by environmental genomics: A review of current strategies with an implementation roadmap

Tristan Cordier1 | Laura Alonso-Sáez2 | Laure Apothéloz-Perret-Gentil1 | Eva Aylagas3 | David A. Bohan4 | Agnès Bouchez5 | Anthony Chariton6 | Simon Cree7 | Larissa Frühe8 | François Keck5 | Nigel Keeley9 | Olivier Laroche9 | Florian Leese10,11 | Xavier Pochon12,13 | Thorsten Stoeck8 | Jan Pawlowski1,14,15 | Anders Lanzén2,16

1Department of Genetics and Evolution, Science III, University of Geneva, Geneva, Switzerland
2AZTI, Marine Research, Basque Research and Technology Alliance (BRTA), Spain
3Red Sea Research Center (RSRC), Biological and Environmental Sciences and Engineering (BESE), King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia
4Agroécologie, INRAE, University of Bourgogne, University Bourgogne Franche-Comté, Dijon, France
5UMR CARTEL, INRA, USMB, Thonon-les-Bains, France
6Department of Biological Sciences, Macquarie University, Sydney, NSW, Australia
7School of Natural Sciences, Bangor University, Gwynedd, UK
8Department of Ecology, Technische Universität Kaiserslautern, Kaiserslautern, Germany
9Benthic Resources and Processes Group, Institute of Marine Research, Tromsø, Norway
10Aquatic Ecosystem Research, Faculty of Biology, University of Duisburg-Essen, Essen, Germany
11Centre for Water and Environmental Research (ZWU), University of Duisburg-Essen, Essen, Germany
12Coastal & Freshwater Group, Cawthron Institute, Nelson, New Zealand
13Institute of Marine Science, University of Auckland, Warkworth, New Zealand
14ID-Gene Ecodiagnostics, Geneva, Switzerland
15Institute of Oceanology, Polish Academy of Sciences, Sopot, Poland
16Basque Foundation for Science, IKERBASQUE, Bilbao, Spain

Correspondence
Tristan Cordier, Department of Genetics and Evolution, University of Geneva, Science III, Geneva, Switzerland.
Email: tristan.cordier@gmail.com

Funding information
Swiss National Science Foundation, Grant/Award Number: 31003A_179125; European Cross-Border Cooperation Programme; 'Ramón y Cajal' contract, Grant/Award Number: RYC-2012-11404; Spanish Ministry of Economy and Competitiveness; Saudi Aramco-KAUST Center for Marine Environmental Observations; French Agence

Abstract
A decade after environmental scientists integrated high-throughput sequencing technologies in their toolbox, the genomics-based monitoring of anthropogenic impacts on the biodiversity and functioning of ecosystems is yet to be implemented by regulatory frameworks. Despite the broadly acknowledged potential of environmental genomics to this end, technical limitations and conceptual issues still stand in the way of its broad application by end-users. In addition, the multiplicity of potential implementation strategies may contribute to a perception that the routine application of this methodology is premature or "in development", hence restraining regulators from binding these tools into legal frameworks. Here, we review recent implementations of...
1 | THE NEED FOR BROAD SCALE ECOSYSTEM MONITORING STRATEGIES

Biodiversity drives the fundamental processes of ecosystems and provides invaluable services on which we depend. Anthropogenic, detrimental impacts on ecosystems, including accelerating climate change, are unprecedented (Waters et al., 2016) and have led to a decline of biodiversity across the globe (Butchart et al., 2010; Cardinale et al., 2012; Hughes et al., 2018). Recent reports stress that one out of the 8 million known species are presently at risk of extinction (IPBES, 2019). This threatens ecosystem function(ing) and services. Therefore, the urgent challenge is now to build a set of efficient tools to enhance our capacity to predict or detect early warnings of critical ecological shifts efficiently, in order to forecast the direction of such shifts and their impacts on ecosystem functions and services (Barnosky et al., 2012; Carpenter et al., 2011; Ratajczak et al., 2018).

Because our societies aim to reach a trade-off between socio-economic development and ecosystems sustainability (UN A/RES/70/1, 2015), regulatory frameworks have been established worldwide for the sustainable development of industries within environmental constraints (Niemeijer, 2002; Grizzetti, Lanzanova, Liqueute, Reynaud, & Cardoso, 2015). Such regulatory systems have been incorporated into various national and international directives, especially for aquatic ecosystems (e.g., the Water Framework Directive, WFD, Directive 2000/60/EC and Marine Strategy Framework Directive, MSFD, Directive 2008/56/EC in Europe, the Clean Water Act of the US Environmental Protection Agency in the USA, as well as the United Nations Convention on the Law of the Sea, UNCLLOS). The backbone of such monitoring programmes is the biological component of ecosystems, as a measure of ecosystem “health” or “integrity” (Karr, 1999). This biological component is often referred to as the Biological Quality Elements in those regulations (BQEs, Borja et al., 2013; Hering et al., 2018). Most monitoring strategies implemented in regulations rely on the bioindication principle (autecology, Box 1), i.e., significant correlations between the occurrence of specific organisms and a set of environmental variables. Although chemical and hydrological monitoring techniques provide an environmental quality snapshot, biological indicators convey a cumulative time-integrated measure as their occurrence is the product of their local adaptation and their responses to ecosystem variations and/or disturbances across an extended period of time (Birk et al., 2012; Carignan & Villard, 2002; Lear, Dopheide, Ancion, & Lewis, 2011).

2 | THE LIMITS OF CURRENTLY IMPLEMENTED ECOSYSTEM MONITORING STRATEGIES

Traditionally, morphologically distinguishable invertebrates have been used as bioindicators in both aquatic and terrestrial ecosystems (Bongers & Ferris, 1999; Gerlach, Samways, & Pryke, 2013; Hodkinson & Jackson, 2005; Reynolds & Metcalfe-Smith, 1992). Fishes, amphibians, macrophytes, phytoplankton and diatoms are also routinely used in aquatic ecosystems (Birk et al., 2012). Various biotic indices (BIs) have been formalized, based on the predictable responses of bioindicator species to environmental disturbances (autecological value) in marine (Borja, Franco, & Pérez, 2000; Maurer, Nguyen, Robertson, & Gerlinger, 1999; Rygg & Norling, 2013), freshwater (Kelly, Perry, & Whiton, 1995; Prygiel & Coste, 2000; Stark, 1998) and terrestrial (Marull, Pino, Mallarach, & Cordobilla, 2007; Urzelai, Hernández, & Pastor, 2000) ecosystems. Almost half of the monitoring methodologies currently used in Europe rely on such BIs (Birk et al., 2012). However, for environments...
BOX 1  Glossary of terms used in this paper

- Implementation strategy: Refers to the way environmental genomics data is produced and analysed in an ecosystem monitoring context. It includes the choice of all the molecular biology steps, i.e., targeted molecules (DNA vs. RNA), metatranscriptomics or metatranscriptomics (shotgun sequencing), and the computational biology steps (analytical approach), i.e., focusing on the taxonomically assigned sequences or considering all the sequences, the use of compositional turnovers (beta-diversity), structural metrics (alpha or phylogenetic diversity and ecological network properties) or functional metrics (functional genes or transcripts diversity).
- Environmental genomics: Suite of molecular tools to sample, process and analyse nucleic acids from an environmental sample (soil, water, sediment, faeces).
- Environmental DNA/RNA: Nucleic acids present in an environmental sample. It encompasses the DNA/RNA within living multi- or unicellular organisms, dead or decaying as well as extracellular material.
- Metabarcoding: A molecular workflow to simultaneously study the diversity of PCR-selected organisms from environmental samples using high-throughput sequencing. This is equivalent to amplicon sequencing of a taxonomic marker.
- Metagenomics: Shotgun sequencing of the genomic DNA isolated from an environmental sample. There is no PCR selection of particular taxonomic group and include coding as well as non-coding genomic material.
- Metatranscriptomics: Shotgun sequencing of retrotranscribed RNA isolated from an environmental sample. As for metagenomics, there is no PCR selection but includes only transcribed RNA (mRNA, rRNA), supposedly functional.
- Bioindicator: A taxon, marker sequence, gene or transcript that is used as an indicator of the ecological status of an environment.
- Ecological network: Representation of statistically inferred biotic interactions through spatial or temporal co-occurrence or co-exclusion. taxa (nodes) are connected by pairwise links (edges). Network ecology aims to understand how these network properties are linked to the functioning of ecosystems.
- Ecological network: Representation of statistically inferred biotic interactions through spatial or temporal co-occurrence or co-exclusion. taxa (nodes) are connected by pairwise links (edges). Network ecology aims to understand how these network properties are linked to the functioning of ecosystems.

3 | ENVIRONMENTAL GENOMICS REVOLUTION FOR BIODIVERSITY RESEARCH AND ECOSYSTEM MONITORING

Over the last decade, the development of environmental genomics (EG) coupled with high-throughput sequencing (HTS) technologies has led to a marked improvement in our ability to document biodiversity patterns, for both species occurrence (amplicon sequencing, i.e., metabarcoding, reviewed in Bohmann et al., 2014; Cristescu & Hebert, 2018; Deiner et al., 2017; Ruppert, Kline, & Rahman, 2019; Taberlet, Bonin, Zinger, & Coissac, 2018; Valentini et al., 2016) and their metabolic functions (metagenomics and metatranscriptomics, reviewed in Ungerer, Johnson, & Herman, 2008; Escalas et al., 2019; Quince, Walker, Simpson, Loman, & Segata, 2017; Singer, Wagner, & Woyke, 2017; Vandenkooijhuyse et al., 2019). Multidisciplinary teams and consortiums have initiated large-scale projects aiming at collecting biodiversity data using EG throughout the globe, to address fundamental ecological questions. Among these initiatives, the large barcoding projects led by the international Barcode of Life (Ribasingham & Hebert, 2007), the Earth Microbiome Project (Gilbert et al., 2010) and the TARA Oceans Project (Karsenti et al., 2011) represent three of the most emblematic examples. Those projects have unravelled an unexpected cryptic (Bickford et al., 2007) and novel microbial diversity (the “unseen majority”) guiding reconstruction of the eukaryotic tree of life (Adl et al., 2019).

Even though this microbial diversity is known to represent a key component of ecosystem functioning (Cavicchioli et al., 2019;
Delgado-Baquerizo et al., 2016; Guidi et al., 2016), the ecology of most microorganisms remains largely enigmatic.

The potential of EG for surveying biodiversity and monitoring natural ecosystems at a broad spatiotemporal scale was quickly identified and implemented by environmental scientists (Baird & Hajibabaei, 2012; Davies et al., 2012; Kelly et al., 2014; Taberlet, Coissac, Pompanon, Brochmann, & Willerslev, 2012). This work has been boosted by the major drop in sequencing costs, with over four orders of magnitude within the last 15 years (https://www.genome.gov). This has enabled numerous clinical and environmental routine applications. Indeed, fueled by the continuous efforts to optimize laboratory protocols and bioinformatic tools, all steps from large-scale collection of samples, generation of HTS data, statistical analysis, and interpretation of results, can now be performed in matter of days or weeks (Deshpande et al., 2019; Juul et al., 2015; Quinn et al., 2016; Reintjes et al., 2019). For aquatic ecosystems especially, the next breakthrough of this revolution is now expected to be the development and deployment of low-cost, automated and miniaturized in situ environmental nucleic acids (eDNA/RNA) samplers (Carr et al., 2017; Gan et al., 2017). These may be integrated into autonomous instruments for broadscale and continuous ecosystem monitoring programmes (Aguzzi et al., 2019; Benway et al., 2019; Bohan et al., 2017; Brandt et al., 2016; Levin et al., 2019).

These advances in genomics-based research have led to a series of pilot studies assessing the applicability of EG for the monitoring of ecosystem changes by collecting biodiversity data from various taxonomic groups (e.g., fishes, macroinvertebrates, protists, bacteria) and environments (e.g., water, biofilms, soil or sediment). Several such pilot studies have targeted multicellular organisms as a replacement for arduous morphological identification of the same taxa (Hajibabaei, Shokralla, Zhou, Singer, & Bair, 2011; Hajibabaei, Spall, Shokralla, & van Konyenburg, 2012; Lejzerowicz et al., 2015; Thomsen et al., 2012; Zhou et al., 2013). However, the potential of EG to leverage the general eukaryotic and prokaryotic diversity for ecological monitoring, has also been explored (Bik Halanych, Sharma, & Thomas, 2012; Chariton, Court, Hartley, Colloff, & Hardy, 2010; Dowle, Pochon, Keeley, & Wood, 2015; Lallias et al., 2015), and indeed advocated (Bouchez et al., 2016; Chariton et al., 2016; Creer et al., 2010; Graham et al., 2016; Payne, 2013). Encouraged by the immense opportunities for ecosystem monitoring, over 45 countries recently decided to join their efforts within the European COST Action DNAqua-Net, to anticipate upcoming paradigm shifts and develop genomic tools tailored for the monitoring of aquatic ecosystems (http://dnaqua.net, Leese et al., 2016). Similarly, other large-scale collaborative projects were recently launched, including STREAM in Canada (https://stream-dna.com/), Lakes380 in New Zealand (https://lakes380.com/) and NGB in France (http://next-genbiomonitoring.org/), aiming at the unbroiling of EG for ecosystem monitoring.

Multiple pilot and methodological EG studies have highlighted important variation in terms of compliance with current regulatory programmes (reviewed in Hering et al., 2018), leading to the proposition of multiple implementation strategies for current and future ecosystem monitoring programmes. Here, we compare and review the strengths and limitations of these EG-based strategies for ecosystem monitoring. Our objective is to pinpoint the criteria of existing monitoring programmes that could be fulfilled by EG methods as of today, and clarify the work ahead for the monitoring programmes that could benefit from EG in the near future, given continued technological and analytical advancements. To this end, we classify these strategies into four broad categories (Figure 1, Table S1): (a) Taxonomy-based analyses that focus on known bioindicator species, or the identification and enumeration of formally or informally described taxa; (b) De novo bioindicator analyses aiming to identify and utilise novel bioindicators, independent of formal taxonomy; (c) Structural community metrics relying on community structure or inferred ecological networks, where taxa are interchangeable; and (d) Functional community metrics or indicators that focus on protein-coding genes or transcripts instead of taxonomic composition. Based on the specificities of each strategy, their level of maturity and their compatibility with existing regulations (Table 1), we propose an implementation roadmap to integrate EG into ecosystems monitoring programmes and highlight future research needs to be undertaken.

4 | "TAXONOMY-BASED" STRATEGY: SCREENING KNOWN SPECIES AND BIOINDICATORS WITH ENVIRONMENTAL GENOMICS

This strategy relies on the enumeration of known biodiversity from DNA obtained from an environmental sample (e.g., sediment, soil, biofilm, water) or from bulk material prepared from an environmental sample by e.g., elutriation, trapped individuals or biofilm scratching (Figure 1a). This strategy closely fits the conventional, morphology-based monitoring approach, because it primarily aims at reaching a satisfactory level of congruence in terms of both qualitative and quantitative biodiversity inventories. The taxonomy-based strategy is de facto limited to the morphologically characterized fraction of biodiversity for which reference sequences are available in public databases. Hence, approaches using it have usually overlooked meiofaunal or microbial taxa, difficult to identify on the basis of morphological traits, and for most of which the autecology is poorly known (but see Pawlowski, Esling, et al., 2016). The reference databases routinely used by EG studies include for instance the universal but essentially uncurated GenBank nucleotide repository from the National Center for Biotechnology Information (Benson et al., 1999, but see Leray, Knowlton, Ho, Nguyen, & Machida, 2019), or the curated databases BOLD for COI barcodes, primarily from animals (Ratnasingham & Hebert, 2007), SILVA for universal ribosomal markers (Quast et al., 2013), PR2 for protists (Guillou et al., 2013), Diat.barcode for diatoms (Rimet et al., 2016), and Unite for fungi (Nilsson, Sharma, Bhatnagar, Bertilsson, & Terenius, 2019).

Depending on the environment assessed and the taxonomic group considered, the performance of taxonomy-based approaches varies considerably (Hering et al., 2018). Benchmarking studies comparing EG-based and conventional morphology-based taxonomic
inventories (Table S1) have shown mixed degrees of congruence. For the noninvasive detection of fish species from DNA traces in filtered marine water, the rate of success from taxonomy-based monitoring is reported near perfect (Thomsen et al., 2012; Bakker et al., 2017; but see DiBattista et al., 2017). For freshwater macroinvertebrate bulk samples, the rate of species detection varied from 67% (Elbrecht, Vamos, Meissner, Aroviita, & Leese, 2017) to 73–83% (Hajibabaei et al., 2011, 2012). In contrast, for benthic diatoms sampled from biofilms, the congruence of morphological taxonomy and EG-inferred taxonomy, in terms of shared taxa at species level, ranged only from 15–18% (Rivera et al., 2018; Vasselon, Domaizon, et al., 2017) to 28% (Visco et al., 2015). The reported congruence for macroinvertebrates sampled from marine sediments ranged from 20% (Lejzerowicz et al., 2015) up to 60% (Aylagas, Borja, Irigoien, & Rodríguez-Ezepeleta, 2016). Noteworthy, those studies also detected numerous species that were unnoticed in morphological inventories (Elbrecht et al., 2017; Hajibabaei et al., 2011, 2012). Despite these discrepancies, the studies inferring BI values from the detected bioindicators species show very promising results, for both freshwater diatoms (Kelly et al., 2018; Kermarrec et al., 2014; Vasselon, Rimet, et al., 2017; Visco et al., 2015) and macroinvertebrates (Elbrecht et al., 2017) as well as for marine macroinvertebrates (Aylagas et al., 2016; Lejzerowicz et al., 2015). While acknowledging that the congruence for both qualitative and quantitative inventories are not fully satisfactory, these studies have demonstrated that EG tools are still able to detect sufficient bioindicator taxa to infer accurate BI values, even when considering only presence/absence (Elbrecht et al., 2016). The EG methodology has therefore been promoted as a promising tool for fast and cost-effective biodiversity screening for ecosystem monitoring, even while the simultaneous collection of classical morphological samples for validation is univocally suggested. Nonetheless, further improvements in molecular protocols as well as BI intercalibration is a necessity towards harmonization and standardization across Europe (Hering et al., 2018; Poikane et al., 2014) and beyond (Jeunen et al., 2019).

Various biological and technical limitations still impede the implementation of the taxonomy-based strategy for routine monitoring applications (Leese et al., 2018). These limitations mainly stem from the fact that the methods sample fundamentally different units of presence (molecules vs. individuals), resulting in different biases affecting richness, abundance and taxonomic composition. The richness of “molecular species”, i.e., operational taxonomic units (OTUs) or Amplicon Sequence Variants (ASVs, the new operational unit paradigm, Callahan, McMurdie, & Holmes, 2017), should not be considered analogous to morphospecies richness even in the theoretical absence of noise resulting from PCR and sequencing biases. This discrepancy is due to cryptic diversity (Stork, 2018), intragenomic or intraspecific marker variation (Bik, Fournier, Sung, Bergeron, & Thomas, 2013; Sun, Jiang, Wu, & Zhou, 2013), and the presence of DNA from dead and inactive organisms or as extracellular DNA (Collins et al., 2018). Likewise, the abundance of taxa inferred from HTS read counts can typically not be used to infer the number of individuals. Indeed, the number of sampled DNA molecules and
sequence reads are a consequence of the number of individuals, but also of the biomass and the variable number of copies of the targeted marker in the genome (Bik et al., 2013; Větrovský & Baldrian 2013), in addition to variations in DNA extractability and primer-specific amplification bias (Elbrecht & Leese, 2015; Krehenwinkel et al., 2017; Piñol, Mir, Gomez-Polo, & Agustí, 2015). Finally, EG studies suffer from a strong sampling effect because DNA extractions are typically performed from small amounts of material, making large-size organisms less well-represented in eDNA extracts (Lanzén, Lekang, Jonassen, Thompson, & Troedsson, 2017). However, bulk samples (Elbrecht et al., 2017), larger extraction volume (Nascimento, Lallias, Bik, & Creer, 2018) or more aggressive homogenization (Lanzén et al., unpublished data) can partially alleviate this issue.

Since the taxonomy-based strategy depends on reference sequences for organism identification, the incompleteness of reference databases can also have a major impact. Hence, completing databases, both by the “vertical” addition of more taxa and by the “horizontal” coverage of wider geographical areas, would certainly contribute to an improvement in identification (McGee, Robinson, & Hajibabaei, 2019; Vasselon, Domaizon, et al., 2017). However, despite sustained efforts, reference databases will probably remain skewed towards some taxa, while suffering from important gaps across other taxonomic groups or biogeographical regions (McGee et al., 2019; Weigand et al., 2019). All these issues directly impact both of the key parameters for applying BIs to assess impact, namely the qualitative and quantitative measures of biodiversity (Pawlowski et al., 2018).

Nevertheless, multiple studies have shown that there is room for considerable improvements to better bridge the current gaps between taxonomy-dependent molecular and morphology-based approaches. Taxonomic breadth in HTS data could be broadened by carefully designing novel amplification primers (Elbrecht et al., 2019) or using more than one primer pair (Corse et al., 2019). Applying correction factors to read counts, based on established knowledge of the biovolume (Vasselon et al., 2018), the number of copies of the targeted marker (Větrovský & Baldrian 2013) or by spiking samples with known internal standard for quantitative determinations (Ji et al., 2020; Tkacz, Hortalá, & Poole, 2018), are all promising methods for resolving these challenges. Finally, the integration of bioinformatic tools for the automated curation of databases from mislabeled sequences will improve their reliability (Ashelford, Chuzhanova, Fry, Jones, & Weightman, 2005; Kozlov, Zhang, Yilmaz, Glöckner, & Stamatakis, 2016).

5 | “DE NOVO” STRATEGY: DISCOVERING NEW BIOINDICATORS AND HARNESING THEM FOR ROUTINE MONITORING

In contrast to the taxonomy-based strategy, the de novo one does not immediately generate an ecological assessment, because it does not employ previous knowledge associated with bioindicators. Instead, the de novo strategy aims at establishing new bioindicators using EG-based profiling of communities and independently generated ecological status or known disturbance gradients (Figure 1b). Harnessing EG and HTS technologies to explore a broader range of biological diversity, formally labelled or not (i.e., taxonomically described or identified), represents an opportunity to move towards a more holistic monitoring paradigm (Bik et al., 2012; Chariton et al., 2010). By considering all the OTU (or ASV) profiles along a known impact gradient of typical anthropogenic origin, studies applying this strategy have shown that HTS data represent a virtually unlimited reservoir of new bioindicators. Examples (listed in Table S1) include contamination by pesticides (Thompson et al., 2016; Andújar et al., 2017) or other agricultural stressors (Salis, Bruder, Piggott, Summerfield, & Matthaei, 2017), and gradients of eutrophication and urban contamination in freshwater systems (Apothéloz-Perret-Gentil et al., 2017; Martínez-Santos et al., 2018; Simonin et al., 2019; Tapolczai, Vasselon, et al., 2019; Tapolczai, Vasselon, et al., 2019). In marine environments, the utility of this strategy has been demonstrated after an oil spill (Bik et al., 2012), in the vicinity of offshore drilling platforms (Cordier, Frontalini et al., 2019; Lanzén, Lekang, Jonassen, Thompson, & Troedsson, 2016; Laroche, Pochon, et al., 2018; Laroche et al., 2016) and aquaculture sites (Pawlowski, Esling et al., 2014, Pochon et al., 2015; Dowle et al., 2015; Keeley, Wood, & Pochon, 2018; Stoeck, Frühe, et al., 2018; Stoeck, Frühe, et al., 2018 as well as along eutrophication and urban or industrial contamination gradients in estuaries (Chariton et al., 2010; Chariton et al., 2015; Angly et al., 2016; Lallias et al., 2015; Obi et al., 2016). Interestingly, most of the studies sampling marine sediments highlighted that meiofaunal invertebrates, such as nematodes, gastrotrichs and platyhelminths (Bik et al., 2012; Chariton et al., 2010; Lanzén et al., 2016), large groups of protists such as diatoms, oomyctes and ciliates (Lanzén et al., 2016; Stoeck, Frühe, et al., 2018 or foraminifera (Pawlowski, Esling, et al., 2014; Laroche et al., 2016; Frontalini et al., 2018) but also fungi (Bik et al., 2012) and bacteria (Angly et al., 2016; Aylagas et al., 2017; Dowle et al., 2015; Keeley et al. 2018; Martínez-Santos et al., 2018; Obi et al., 2016; Stoeck, Frühe, et al., 2018) have great potential as bioindicators of anthropogenic impacts and can readily be captured by EG studies.

Unfortunately, most proof-of-concept studies employing the de novo strategy have not yet validated their results by performing ecological assessments based on newly identified bioindicators as a reference in a new environmental context. For this information to be useful on new samples, the data obtained from known disturbance gradients (i.e., reference or training data set) must be operational in different spatiotemporal contexts. To this end, two main approaches have been proposed and tested, namely indicator value (e.g., the IndVal approach, Dufrêne & Legendre, 1997) and supervised machine learning (SML, Crisci, Ghattas, & Perera, 2012; Libbrecht & Noble, 2015).

The indicator value approach ascribes autecological values (or discrete “eco-groups”) to OTUs or ASVs based on their occurrence in samples of known disturbance level, in a similar manner as for the establishment of morphology-based bioindicators. Hence, the autecological values of these de novo bioindicators are directly calibrated on the HTS data, which alleviates the qualitative
and quantitative biases encountered with the taxonomy-based EG strategy. This has proven successful for both freshwater benthic diatoms (Apothéloz-Perret-Gentil et al., 2017; Tapolczai, Vasselon, et al., 2019; Tapolczai, Vasselon, et al., 2019) and for bacterial and eukaryotic communities in streams and estuarine systems (Chariton et al., 2015; Li et al., 2018). An analogous approach is the use of polynomial quantile regression splines (Anderson 2008). This has shown great promise for the prediction of impacts from organic enrichment in aquaculture sites using eukaryotic and prokaryotic metabarcoding data in parallel (Keeley et al., 2018). For diatoms, the accuracy of the assessment can be largely improved, arguably because the indicator value approach makes use of a larger number of OTUs or ASVs, compared to an approach relying solely on their taxonomic assignments (Apothéloz-Perret-Gentil et al., 2017; Tapolczai, Vasselon, et al., 2019; Tapolczai, Vasselon, et al., 2019).

Supervised machine learning (SML) also requires training data sets, i.e., reference disturbance levels (labels) associated with the community profiles of the samples (features). These algorithms are best at classification problems involving multidimensional and noisy data sets (Libbrecht & Noble, 2015), which are common attributes

Figure 1. Overview of the current methodology for the monitoring of ecosystems, that relies mostly on the morphological identification of biodiversity and/or bioindicators of anthropogenic impacts. Ecological diagnostics are performed based on reference biodiversity or on reference biotic indices for a given ecosystem. The development of environmental genomics methodologies has led to the proposition of multiple implementation strategies that can intervene at different levels of the monitoring workflow, to produce an ecological diagnostic. Green colours and smileys within boxes indicate reference biodiversity and "good" or "high" ecological status while red colours and smileys represent nonreference biodiversity and "poor" ecological status (i.e., impacted environments) Green colours and smileys within boxes indicate reference biodiversity and "good" or "high" ecological status while red colours and smileys represent nonreference biodiversity and "poor" ecological status (i.e., impacted environments). The colours on tags besides organisms or sequences indicate their bioindication value (red: indicator of impact, yellow: indicator of intermediate status, green: indicator of good status). In this review paper, these strategies have been grouped into four broad categories: (a) Taxonomy-based analyses focused on identification of known bioindicators or described taxa; (b) De novo bioindicator analyses; (c) Structural community metrics including inferred ecological networks; and (d) Functional community metrics (metagenomics or metatranscriptomics).
of HTS data. The task is to automatically disentangle the feature signal (OTU or ASV profiles) and their co-occurrence that convey an ecological signal from background noise. This extracted knowledge is self-contained in a trained model that can be used to make predictions of disturbance level on new samples, based on their compositional profiles (Cordier, Lanzén, et al., 2019). Supervised machine learning also alleviates the qualitative and quantitative biases that hamper the taxonomy-based strategy in a more straightforward manner, because the model is trained directly on HTS data. The applicability of SML has been demonstrated in marine environments, for the detection of various pollutants (Smith et al., 2015) and for the prediction of aquaculture impacts on benthic biodiversity (Cordier et al., 2017, 2018). The SML-based inference of BI values has also been shown to outperform the taxonomy-based strategy, relying on the detection of established macroinvertebrates bioindicators DNA (Cordier et al., 2018), and may be more powerful than the IndVal approach (Frühe et al., 2020). Supervised machine learning applications have also succeeded in predicting the origin of container ship ballast waters (Gerhard & Gunsch, 2019).

The de novo strategy provides numerous advantages over the taxonomy-based one. First, it can reduce or bypass the dependence on reference sequence databases for taxonomic assignments of HTS reads to known bioindicators. Instead, new ecological knowledge is hypothesised de novo during the calibration of OTUs or ASVs attributes or during the supervised training of a model (SML). Second, it can leverage powerful but previously inaccessible groups of bioindicators among prokaryotes, protists, meiofauna and mesozooplankton, that are widespread and may react both faster and stronger to environmental disturbances (Bouchez et al., 2016; Creer et al., 2010; Pawlowski, Esling, et al., 2016; Payne, 2013). Finally, when applied for the inference of BIs that are currently employed in routine monitoring programmes, a de novo strategy is directly compatible with current regulations, because the assessment categories remain the same and the BI values are simply inferred indirectly. Hence, this strategy assures a full backward and forward compatibility with current monitoring programmes, facilitating continuity of important time series data sets ( Bálint et al., 2018).

6 | "STRUCTURAL COMMUNITY METRICS" STRATEGY: BLENDING THEORETICAL ECOLOGY INTO ROUTINE ECOSYSTEM MONITORING

This strategy relies on metrics extracted from the community structure or from inferred ecological networks – where taxa are interchangeable – in order to assess the impact of disturbance and its ramifications on ecosystem functioning (Figure 1c). This represents a clear paradigm shift for ecosystem monitoring programmes, because the evaluation of bioindicators, based on the compositional variation of communities, is not the main aim of the strategy. Instead, its focus is to discover and understand the ecological processes shaping biological communities and their response to disturbances, which is indeed one of the core questions of ecological research. It has long driven the exploration of the links between generic, taxonomy and composition-independent biodiversity metrics or species functional traits distribution and ecosystems functioning and resilience, to reach a more general theoretical framework (Cardinale, Nelson, & Palmer, 2000; Hooper et al., 2005; Ives & Carpenter, 2007; Loreau & de Mazancourt, 2013; McCann, 2000; Mouillot, Graham, Villéger, Mason, & Bellwood, 2013; Tilman, Reich, & Knops, 2006).

Structural community metrics can be computed from compositional data generated by EG studies, including alpha diversity (e.g., OTU or ASV richness, Shannon diversity or Pielou evenness; reviewed in Daly, Baetens, & De Baets, 2018), along with its phylogeny-aware derivatives (reviewed in Tucker et al., 2017; Washburne et al., 2018). Under anthropogenic impact, alpha diversity in marine sediment has been found to decrease for foraminifera (Laroche, Wood, et al., 2018; Pawlowski, Esling, et al., 2014, 2016), ciliates (Stoeck, Kochems et al., 2018) and bacterial communities (Stoeck, Frühe, et al., 2018). Conversely, disturbances in marine sediments can also trigger increases in bacterial diversity and metabolic activity (Galand et al., 2016; Pérez-Valera et al., 2017). This suggests that the variation of alpha diversity alone is insufficient as a widely applicable indicator of disturbance. Phylogeny-aware metrics attempt to account for the evolutionary relationships among taxa composing communities, to provide insights into community assembly processes and by extension their predictable responses to environmental variations (Webb, Ackerly, McPeek, & Donoghue, 2002; Cavender-Bares, Kozak, Fine, & Kembel, 2009, but see Mayfield & Levine, 2010; Gerhold, Cahill, Winter, Bartish, & Przing, 2015). This relationship between phylogenetic diversity and ecosystem functioning has received a lot of attention by plant ecologists (Flynn, Mirochnick, Jain, Palmer, & Naeem, 2011). However, only few studies have employed EG data to this end, targeting mostly microbial groups, which, as for simple alpha-diversity metrics, has resulted in contrasting conclusions (Galand, Salter, & Kalenitchenko, 2015; Pérez-Valera et al., 2017; Liu et al., 2017; but see Venail & Vives, 2013; Keck & Kahler, 2019 for studies employing sequencing data but not strictly EG).

Metrics based upon alpha diversity may be misleading (Santini et al., 2017) because their variation is often nonlinear, strongly scale-dependent (Chase et al., 2019) and valuable only in comparing contexts sampled using the same methodology (Shade, 2017). It also implicitly conveys the idea that “higher diversity is better” which is not necessarily true (Shade, 2017). The inference of ecological functioning based on phylogeny-aware metrics relies on the niche conservatism concept, which postulates that closely related taxa share similar functional traits (Cavender-Bares et al., 2009; Srivastava, Cadotte, Macdonald, Marushia, & Mirochnick, 2012; Webb et al., 2002). Under this assumption, increased phylogenetic diversity may support functionally diverse or multifunctional ecosystems (Hector & Baghi, 2007 but see Manning et al., 2018). By extension, higher phylogenetic diversity may also support ecosystem resilience, provided that the species fulfilling similar functions have differing responses to disturbances (Cadotte, Dinnage, & Tilman, 2012; Oliver et al., 2015). However, because not all functional traits
necessarily have a phylogenetic signal (Srivastava et al., 2012), including for microbes (Martin, Treseder, & Pusch, 2013), inferring ecosystem functioning and the level of anthropogenic impact based on phylogeny-aware metrics alone may prove to be misguided. Likewise, conservation strategies based on these metrics may also be suboptimal (Mazel et al., 2018).

Another set of structural community metrics can be computed from the topology of inferred ecological or co-occurrence networks, representing potential biotic interactions (reviewed in Faust & Raes, 2012; Layeghfard et al., 2017; Vacher et al., 2016). Based on empirical evidence of the variation in network structure under environmental disturbance (Karimi, Meyer, Gilbert, & Bernard, 2016; Ma et al., 2019; Tylanakis, Tscharrmke, & Lewis, 2007; Zhou et al., 2011), their properties have been suggested as potential indicators of ecosystem functioning and integrity (Bohan, Caron-Lormier, Mugler, Raybould, & Tamaddoni-Nezhad, 2011; Bohan et al., 2017; Delmas et al., 2019; Gray et al., 2014; Karimi et al., 2017; Lau, Borrett, Baiser, Gotelli, & Ellison, 2017; Pellissier et al., 2018; Tylanakis & Morris, 2017). In recent years, a growing interest in these approaches has led to a series of studies employing EG to infer ecological networks from microbial community data (Lupatini et al., 2014; Pauvert, Vallance, Delière, Buée, & Vacher, 2019; Pérez-Valera et al., 2017; Zappelini et al., 2015; Zhou et al., 2011) or from macroinvertebrates (Compson et al., 2019), in order to explore the links between network properties such as connectance, centrality or nestedness, and ecosystem functioning. For instance, it has been shown that bacterial communities in anthropized soil may have fewer potentially interacting taxa, than in natural soil (Lupatini et al., 2014). Likewise, in aquatic ecosystems, anthropogenic impacts are reflected in co-occurrence networks by a lower connectivity (Laroche, Wood, et al., 2018; Lawes, Dafforn, Clark, Brown, & Johnson, 2017; Li et al., 2018) and a lower ratio of positive interactions (Laroche, Wood, et al., 2018).

While promising, exploring the links between the properties of ecological networks inferred from EG data and ecosystem functioning is still in its infancy (Faust et al., 2012, 2015; Laroche, Wood, et al., 2018; Lawes et al., 2017; Li et al., 2018; Lima-Mendez et al., 2015; Pauvert et al., 2019). Multiple methodological issues limit the inference of robust networks from EG data based on co-occurrences in space or time. For example, read counts are strictly compositional, representing relative abundance of the marker itself, rather than presence or absolute abundances (but see Friedman & Alm, 2012; Kurtz et al., 2015). Further, it is challenging to control for covariates and confounding environmental parameters (but see Chiquet, Mariadassou, & Robin, 2018; Cougoul, Bailly, & Wit, 2019; Momal, Robin, & Ambroise, 2019; Tackmann, Matias Rodrigues, & von Mering, 2019; Tamaddoni-Nezhad et al., 2013), replicability of inference (Pauvert et al., 2019) and the relative merits of statistical and logical inference (Vacher et al., 2016). Robust networks also require considerably more replicates than are typically collected in EG studies, which increase both time and costs. Nevertheless, as more benchmark data sets containing both EG data and independently confirmed interactions between taxa become available to complement simulated data sets (see Lima-Mendez et al., 2015), making robust simulated data sets to explore the applicability of their metrics for ecosystem monitoring will probably come within reach in the years to come.

7 | "FUNCTIONAL COMMUNITY METRICS" STRATEGY: EMPLOYING FUNCTIONAL ENVIRONMENTAL GENOMICS FOR ROUTINE MONITORING

Another avenue of implementation of EG for ecosystem monitoring is the use of shotgun metagenomics and metatranscriptomics, depicting the metabolic capabilities of the community, and the expressed genes at the moment of sampling, respectively (Figure 1d). However, ecologists have yet to disentangle the relative importance and relationship of taxonomic diversity and functional traits for ecosystems functioning (Flynn et al., 2011; Gagic et al., 2015). This is particularly true in microbial ecology with the "who's there" versus "what they are doing" paradigms that often relate to the employed molecular methodologies, i.e., metabarcoding versus metagenomics and metatranscriptomics (Xu, Malmer, Langille, Way, & Knight, 2014). Some metagenomic contigs and functional transcripts were indeed found to represent efficient bioindicators of anthropogenic disturbances (Table S1), in terrestrial (de Menezes, Clipson, & Doyle, 2012), groundwater (He et al., 2018), freshwater (Cheaib, Boulch, Le Mercier, & Derome, 2018; Falk et al., 2019; Thompson et al., 2016) and marine environments (Birrer et al., 2019; Galand et al., 2016; Kisand, Valente, Lahm, Tanet, & Lettieri, 2012), opening up potential avenues for future routine ecosystem monitoring applications. Functional and taxonomic profiles may respond differently under anthropogenic disturbance (Cheaib et al., 2018), as well as under natural environmental variation (Barberán, Fernández-Guerra, Bohannan, & Casamayor, 2012; Louca, Jacques, et al., 2016; Louca et al., 2018). This taxon-function decoupling paves the way towards a molecular trait-based ecology (Lajoie & Kembel, 2019; Raes, Letunic, Yamada, Jensen, & Bork, 2011).

In an ecosystem monitoring context, functional profiles present two important features that anticipate these proxies to be more accurate than taxonomic profiles for the detection of a given environmental disturbance. First, because prokaryotes functional redundancy may be widespread (Louca et al., 2018; Pearlman et al., 2019; but see Galand, Pereira, Hochart, Auguet, & Debros, 2018 and Ramond et al. 2019 for protists), any given anthropogenic disturbance might trigger a similar response across multiple taxonomic groups. Under this assumption, ecosystem monitoring based on functional profiles may be less sensitive to biogeographical effects, random demographic drift, and species dispersal limitation than a monitoring strategy based on taxonomic profiles. This functional redundancy would also allow the establishment of a direct and mechanistic link between a measured functional response to a given anthropogenic disturbance. Second, because functional shifts are likely to occur prior to compositional ones, as a response of the taxa present to
the disturbance, the variation of functional profiles may constitute useful early warnings for a timelier ecosystem management, especially the ones detected by means of metatranscriptomics. However, RNA molecules are reportedly less stable than genomic DNA, which would add challenging practical constraints that could preclude their implementation in routine ecosystem monitoring programmes (but see von Ammon et al., 2019; Cristescu, 2019; Fordyce et al., 2013; Pochon, Zaiko, Fletcher, Laroche, & Wood, 2017). As a possible cost-effective “shortcut”, bacterial 16S rRNA profiles can be used to predict functional community profiles, based on evolutionary models (Aßhauer, Borja, Irigoien, & Rodríguez-Ezpeleta, 2015; Langille et al., 2013). Thus, 16S data could be also explored for searching potential functional bioindicators by this approach (Cordier, 2020; Laroche, Wood, et al., 2018; Mukherjee et al., 2017).

8.1 | The emergence of standards for EG methodologies to be applied for monitoring programmes

The time lag between technological breakthroughs, the uptake by scientists and the implementation of research results into real management applications can be notoriously long. Even for clinical applications where the contributions of genomics have long been anticipated (Dulbecco, 1986; Manolio et al., 2013) and for which economic perspectives are obvious, its implementation for routine healthcare applications is considered to have started five years ago (Stark et al., 2019). This is three times faster than the average 17 years for any healthcare research (Morris, Wooding, & Grant, 2011). The emergence of consensus standards for methodological protocols and data formats for interoperable exchanges, represent the most challenging issue for the routine adoption (Stark et al., 2019).

The field of EG for ecosystems monitoring is experiencing similar issues and has yet to overcome some of the barriers to the necessary paradigm-shift in monitoring programmes (Hering et al., 2018). Some of the noteworthy steps towards this goal were achieved with the widespread adoption of the MIGS, MIMARKS and MIXS standards in genomics, specifying the minimum information that should accompany any genome, marker gene sequences or any sequence (Field et al., 2008; Yilmaz et al., 2011). Now the most challenging part resides in the adoption of standardized methodologies to produce, store and analyse EG data for a given environmental setting. Given the variety of biological models and environmental matrices, reaching a consensus in the scientific community and formalizing standards appears very challenging, especially for metabarcoding (Knight et al., 2016; Pollock, Glendinning, Wisedchanwet, & Watson, 2018; Wilcox et al., 2018; Zinger et al., 2019) and its application to ecosystem monitoring (Cristescu & Hebert, 2018; Hering et al., 2018). Yet, these hurdles are not specific to genomics methodologies, but also exist for the morphology-based ones (Birk et al., 2012). Building robust, shared methodological standards is of course necessary and important efforts are deployed to reach this aim (Hering et al., 2018; Leese et al., 2018; Working Group CEN/TC230/WG28), for the sampling of eDNA (CEN, 2018a; Dickie et al., 2018; Wilcox et al., 2018), the molecular protocols (Blackman et al., 2019; Goldberg et al., 2016) as well as for bioinformatics (Knight et al., 2018; Roy et al., 2018), data interoperability (Callahan et al., 2017; McDonald et al., 2012) and reference databases (CEN, 2018b).

8.2 | Matching the right implementation strategy to the right monitoring programme

Several monitoring programmes may benefit quickly and reliably from an EG implementation, while others may require further optimization of molecular protocols or adjustments of their assessment criteria (Table 1). For instance, monitoring programmes relying primarily on taxonomic inventories are still hindered by the lack of congruence between the recovered species list and their relative abundances, even though the biological and technical biases might be partially alleviated in the future. Furthermore, despite the sustained effort, reference sequence databases for barcoding remain skewed toward some groups and geographical locations (McGee et al., 2019; Weigand et al., 2019), limiting congruence between EG and morphotaxonomic inventories. Hence, the taxonomy-based implementation strategy for these monitoring programmes will require improvements of molecular protocols and reference databases, to generate EG data that better fit the current standards, or an adaptation of the currently implemented assessment criteria to fit the specificities of EG data (Hering et al., 2018).

Monitoring programmes relying on the screening of established bioindicators for the computation of BI values are proposed as being compatible with an implementation of EG (Hering et al., 2018; Pawlowski et al., 2018). Indeed, this compatibility is greatly facilitated by the fact that the assessment criteria, i.e., BIs, are not meant to strictly rely on taxonomic inventories but rather on the autecology of bioindicators. Hence, for the taxonomy-based strategy, the BI formulations can compensate the impact of taxonomic mismatches between morphology and EG and databases incompleteness to some extent, because multiple taxa are ascribed identical autecological values, conveying similar ecological signal (Keck, Vasselon, Rimet, Bouchez, & Kahlert, 2018). The applicability of this approach has been demonstrated in freshwater (Elbrecht et al., 2017; Kelly et al., 2018; Mortagau et al., 2019; Rivera, Vasselon, Bouchez, & Rimet, 2020; Vasselon, Domaizon, et al., 2017) and in marine environments (Aylagas et al., 2016; Lejzerowicz et al., 2015). However, those studies have also shown that a large amount of sequences are not taxonomically assigned and currently omitted for ecological assessment, opening the door to new approaches that could extract ecological information from those unlabelled sequences.

The de novo strategy uses the occurrence of previously scrutinized sequences in samples of known BI values or other impact
measures to ascribe autecological values to sequences directly, or generate a predictive model (Apothéloz-Perret-Gentil et al., 2017; Cordier et al., 2017; Tapolczai, Vasselon, et al., 2019). Hence, these approaches are less sensitive to the biological and technical issues mentioned above, because the ecological signal (autecology) is calibrated directly on the specificities of EG data. From an implementation perspective, this de novo strategy thus may represent the most direct path towards implementation of EG into monitoring programmes relying on BIs (Figure 2). Though somewhat unintuitive, this is because inferred BI values with a de novo strategy convey the same ecological meaning as they do with current methodologies, which is not the case when BIs values are inferred from biodicators composition profiles depicted by EG data, as their autecological values were calibrated only on morphology-based data. Thus, the de novo strategy assures a better continuity with previous BIs data and time series and expand the range of possible biodicators to virtually any taxa or sequence.

Structural and functional community metrics represent alternative implementation strategies that may ultimately lead to a more generic, broadly applicable ecological monitoring framework (Bohan et al., 2017; Escalas et al., 2019; Karimi et al., 2017; Pellissier et al., 2018; Quince et al., 2017; Singer et al., 2017; Tylanakis & Morris, 2017). These strategies hold the potential to provide a more mechanistic and functional understanding of the response of biological communities to ecosystem variation. Such knowledge could hence be included in predictive models to forecast shifts in biodiversity structure and possibly their consequences on their associated ecosystem services under different disturbance scenarios. However, an operational ecosystem monitoring framework remains to be built upon this theoretical ecological work (Figure 2), that has only partially been experimentally validated (but see Laroche, Wood, et al., 2018; Ma et al., 2019). In addition, the extraction of structural or functional community metrics remain active fields of ecological research, and the emergence of a molecular trait-based ecology using metagenomics and metatranscriptomics profiles is in its infancy (Lajoie & Kembel, 2019). Hence, it is premature to discuss their operational implementation and regulatory establishment, but their ecological benefit should be anticipated. Nevertheless, the collected labelled data sets including samples for the production of EG data in the course of future ecosystem monitoring campaigns will certainly contribute to move these possibilities forward.

8.3 Collecting reference data and eDNA/eRNA samples in parallel

If EG-based methods are to complement or replace current morphology-based ones, the prerequisite is to establish whether they can provide similar ecological diagnostics, to ensure a smooth implementation and compatibility with existing time series (Bálint et al., 2018; Leese et al., 2016). This inevitably implies extensive parallel sampling of currently implemented and EG methodologies for some time, to build reference data sets on which the applicability can be assessed and the calibration with previous methodology performed (Keeley et al. 2018; Leese et al., 2016). To be reliable, such reference data sets have to cover a broad range of possible environmental conditions for a given ecosystem across multiple spatiotemporal scales, ideally in a balanced manner, to account for biotic interactions, random demographic drift and dispersal limitations that may interact with the anthropogenic pressures in the assembly of communities.

The collection of reference data raises concerns regarding the substantial financial investment necessary for monitoring programmes adopting one or a combination of EG strategies, versus the “risk” of technological novelty and/or paradigm shift. However, the collected reference data sets would still be extremely valuable in such case, because the extracted DNA/RNA alongside the accompanying reference metadata can be safely stored and reanalysed later on, assuring a forward compatibility to the limit of availability of stored DNA/RNA material (Hering et al., 2018; Jarman, Berry, & Bunce, 2018). Indeed, molecular costs are usually far less prohibitive than those related to field sampling and metadata collection. Hence, such fully labelled data sets will constitute the ideal benchmarks against which to assess the validity of any new implementation strategy based on novel technology or new paradigm.

9 CONCLUSION AND FURTHER RESEARCH NEEDS

The potential for EG-based methods for ecosystems monitoring is enormous and can presently fulfil most of the requirements of current monitoring programmes. Moving towards a routine use of EG is certainly a paradigm shift, but this technological breakthrough will overcome the limitations of current morphotaxonomy methodologies and enable the required upscaling to meet monitoring needs in a changing world. Without doubts, EG-based methods will pave the way for a more cost-effective, faster, reproducible and semi-automatable ecosystem monitoring framework. Regardless of the implementation strategy envisioned, the following key technological, scientific and societal improvements will be beneficial for a smoother transition:

- A collaborative and transdisciplinary design of monitoring campaigns, involving both experts, stakeholders and regulators would allow monitoring programmes to more easily bridge the science-policy gap.
- A collection of reference morphological and molecular data in parallel, at least in a subset of reference points or during a transition period, will assure backward and forward compatibility of time series data sets, regardless of the envisioned implementation strategy to be decided in future monitoring campaigns.
- The efforts to complete reference sequence databases need to be sustained, by adding more representatives of the known biodiversity, with a wider geographical coverage.
- A reference database framework for de novo strategies needs to be established. A key requirement is the ability to reliably compare
FIGURE 2 Strengths and limitations of the currently envisioned implementation strategies of environmental genomics for the monitoring of ecosystems, and their ability to fulfill the criteria of existing monitoring programmes. The tag above organisms indicate their taxonomic affiliation and smileys indicate the ecological status of the sample (from "poor" to "good")

- OTUs or ASVs identified in monitoring programmes to formally establish knowledge about their sensitivity to disturbance.
- The taxonomic resolution level (haplotype, species, genus, family, order, class) at which HTS reads are most informative as genetic bioindicators for a given situation remains to be identified.
- For the identification of novel genetic bioindicators in complex communities, it will be important to distinguish the effect of natural (seasonal) variation from disturbance-induced community changes with rigorous experimental designs.
- Basic and replicable research is highly needed to develop a structural and functional community metrics-based implementation strategy. Such effort will probably contribute to the establishment of a more broadly applicable monitoring framework and less constrained by the database and geographical coverage limitations.

ACKNOWLEDGEMENTS

T.C., J.P., and L.A.P.G. were supported by the Swiss National Science Foundation (grant 31003A_179125). T.C., J.P., L.A.P.G., and A.B. were supported by the European Cross-Border Cooperation Programme (Interreg France-Switzerland 2014-2020, SYNAQUA project). L.A.S. was funded by a 'Ramón y Cajal' contract (RYC-2012-11404) from the Spanish Ministry of Economy and Competitiveness. E.A. is funded by the Saudi Aramo-KAUST Center for Marine Environmental Observations. D.A.B. would like to acknowledge the financial support of the French Agence Nationale de la Recherche project NG8 (ANR-17-CE32-011) and the ERA-NET C-IPM BioAWARE. A.B., and F.K. were supported by the Office Français de la Biodiversité (OFB). S.C. benefitted from the Agence Nationale de la Recherche project NGB (ANR-17-CE32-011). It was supported by the National Natural Science Foundation (grant 31670191) and the Special Funds for Major Scientific Research Projects of the Ministry of Education of China (grant 2015ZD001). It was supported by the National Natural Science Foundation of China (grant 31470132) and the Key Research and Development Program of Shanghai (grant 18511121700).

REFERENCES

Some references are cited in Table S1.

Adl, S. M., Bass, D., Lane, C. E., Lukeš, J., Schoch, C. L., Smirnov, A., ... Zhang, Q. (2019). Revisions to the classification, nomenclature, and diversity of Eukaryotes. *Journal of Eukaryotic Microbiology*, 66(1), 4–119. https://doi.org/10.1111/jeu.12691

Aguzzi, J., Chatzievangelou, D., Marini, S., Fanelli, E., Danovaro, R., Flögel, S., ... Company, J. B. (2019). New high-tech flexible networks for the monitoring of deep-sea ecosystems. *Environmental Science and Technology*, 53(12), 6616–6631. https://doi.org/10.1021/acs.est.9b00409

Anderson, M. J. (2008). Animal-sediment relationships re-visited: Characterising species' distributions along an environmental gradient using canonical analysis and quantile regression splines. *Journal of Experimental Marine Biology and Ecology*, 366(1-2), 16–27. https://doi.org/10.1016/j.jembe.2008.07.006

Andújar, C., Arribas, P., Gray, C., Bruce, C., Woodward, G., Yu, D. W., & Vogler, A. P. (2018). Metabarcoding of freshwater invertebrates to detect the effects of a pesticide spill. *Molecular Ecology*, 27(1), 146–166. https://doi.org/10.1111/mec.14410
Borja, A., Elliott, M., Andersen, J. H., Cardoso, A. C., Carstensen, J., Ferreira, J. G., ... Zampounkas, N. (2013). Good environmental status of marine ecosystems: What is it and do we know when we have attained it? Marine Pollution Bulletin, 76(1–2), 16–27. https://doi.org/10.1016/j.marpolbul.2013.08.042

Borja, A., Franco, J., & Pérez, V. (2000). A marine biotic index to establish the ecological quality of soft-bottom benthos within European estuarine and coastal environments. Marine Pollution Bulletin, 40(12), 1100–1114. https://doi.org/10.1006/mppb.2000.0061

Borrrell, Y. J., Miralles, L., Do Huu, H., Mohammed-Geba, K., & García-Vazquez, E. (2017). DNA in a bottle - Rapid metabarcoding survey for early alerts of invasive species in ports. PLoS One, 12(9), e0183347. https://doi.org/10.1371/journal.pone.0183347

Bouchez, T., Blieux, A. L., Dequiedt, S., Domaizon, I., Dufresne, A., Ferreira, S., ... Ranjard, L. (2016). Molecular microbiology methods for environmental diagnosis. Environmental Chemistry Letters, 14(4), 423–441. https://doi.org/10.1007/s10311-016-0581-3

Brandt, A., Gutt, J., Hildebrandt, M., Pawlowski, J., Schwendinger, J., Soltwedel, T., & Thomsen, L. (2016). Cutting the umbilical: New technological perspectives in benthic deep-sea research. Journal of Marine Science and Engineering, 4(2), 36. https://doi.org/10.3390/jmse4020036

Bucklin, A., Yeh, H. D., Questel, J. M., Richardson, D. E., Reese, B., Brandt, A., Gutt, J., Hildebrandt, M., Pawlowski, J., Schwendner, J., ... Matthews, R. A. (2014). A molecular-based approach for examining responses of eukaryotes in microcosms to contaminant-spiked estuarine sediments. Environmental Toxicology and Chemistry, 33(2), 359–369. https://doi.org/10.1002/etc.2450

Chariton, A. A., Stephenson, S., Morgan, M. J., Steven, A. D. L., Colloff, M. J., Court, L. N., & Hardy, C. M. (2015). Metabarcoding of benthic eukaryote communities predicts the ecological condition of estuaries. Environmental Pollution, 203, 165–174. https://doi.org/10.1016/j.envpol.2015.03.047

Chariton, A. A., Sun, M., Gibson, J., Webb, J. A., Leung, K. M. Y., Hickey, C. W., & Hose, G. C. (2016). Emergent technologies and analytical approaches for understanding the effects of multiple stressors in aquatic environments. Marine and Freshwater Research, 67(4), 414–428. https://doi.org/10.1071/MF15119

Chase, J. M., McGill, B. J., Thompson, P. L., Antão, L. H., Bates, A. E., Blowes, S. A., ... O’Connor, M. (2019). Species richness change across spatial scales. Oikos, 128(8), 1079–1091. https://doi.org/10.1111/oik.05968

Cheaib, B., Boulch, M. L., Mercier, P. L., & Derome, N. (2018). Taxon-function decoupling as an adaptive signature of Lake Microbial metacommunities under a chronic polymetallic pollution gradient. Frontiers in Microbiology, 9(MAY), 1–17. https://doi.org/10.3389/fmicb.2018.00869

Chiquet, J., Mariadassou, M., & Robin, S. (2018). Variational inference for sparse network reconstruction from count data. In Proceedings of the 36th International Conference on Machine Learning (p. PMLR 97:1162–1171). Retrieved from http://arxiv.org/abs/1806.03120

Coelho, F. J. R. C., Cleary, D. F. R., Costa, R., Ferreira, M., Polónia, A. R. M., Silva, A. M. S., ... Gomes, N. C. M. (2016). Multitaxon activity profiling reveals differential microbial response to reduced seawater pH and oil pollution. Molecular Ecology, 25(18), 4645–4659. https://doi.org/10.1111/mec.13779

Collins, R. A., Wangetstein, O. S., O’Gorman, E. J., Mariani, S., Sims, D. W., & Genner, M. J. (2018). Persistence of environmental DNA in marine systems. Communications Biology, 1(1), 1–11. https://doi.org/10.1038/s42003-018-0192-6

Compson, Z. G., Monk, W. A., Hayden, B., Bush, A., O’Malley, Z., Hajibabaei, M., ... Baird, D. J. (2019). Network-based biomonitoring: Exploring freshwater food webs with stable isotope analysis and DNA metabarcoding. Frontiers in Ecology and Evolution, 7(395). https://doi.org/10.3389/FEVO.2019.00395

Cordier, T. (2020). Bacterial communities’ taxonomic and functional turnovers both accurately predict marine benthic ecological quality status. Environmental DNA, 2, 175–183. https://doi.org/10.1002/edn.5

Cordier, T., Estling, P., Lejzerowicz, F., Visco, J., Ouadahi, A., Martins, C., ... Pawlowski, J. (2017). Predicting the ecological quality status of marine environments from eDNA metabarcoding data using supervised machine learning. Environmental Science and Technology, 51(16), 9118–9126. https://doi.org/10.1021/acs.est.7b01518

Cordier, T., Forster, D., Dufresne, Y., Martins, C. I. M., Stoeck, T., & Pawlowski, J. (2018). Supervised machine learning outperforms taxonomy-based environmental DNA metabarcoding applied to biomonitoring. Molecular Ecology Resources, 18(6), 1381–1391. https://doi.org/10.1111/1755-0998.12926
Manning, P., van der Plas, F., Soliveres, S., Allan, E., Maestre, F. T., Mace, G., ... Fischer, M. (2018). Redefining ecosystem multifunctionality. *Nature Ecology and Evolution*, 2(3), 427–436. https://doi.org/10.1038/s41559-017-0461-7

Manolio, T. A., Chisholm, R. L., Ozenberger, B., Roden, D. M., Williams, M. S., Wilson, R., ... Ginsburg, G. S. (2013). Implementing genomic medicine in the clinic: The future is here. *Genetics in Medicine*, 15(4), 258–267. https://doi.org/10.1038/gim.2012.157

Martínez-Santos, M., Lanzén, A., Paula-Calvo, J., Martín, I., Garbiso, C., & Ruiz-Romera, E. (2018). Treated and untreated wastewater effluents alter river sediment bacterial communities involved in nitrogen and sulphur cycling. *Science of the Total Environment*, 633, 1051–1061. https://doi.org/10.1016/j.scitotenv.2018.03.229

Martiny, A. C., Treseder, K., & Pusch, G. (2013). Phylogenetic conservatism of functional traits in microorganisms. *The ISME Journal*, 7(4), 830–838. https://doi.org/10.1038/isme.2012.160

Marull, J., Pino, J., Mallarach, J. M., & Cordobilla, M. J. (2007). A land suitability index for strategic environmental assessment in metropolitan areas. *Landscape and Urban Planning*, 81(3), 200–212. https://doi.org/10.1016/j.landurbplan.2006.11.005

Mauer, D. O. N., Nguyen, H. A. L., Robertson, G., & Gerlinger, T. O. M. (1999). The Infaunal Trophic Index (ITI): Its suitability for marine environmental monitoring. *Ecological Applications*, 9(June), 699–713. https://doi.org/10.1890/1051-0761(1999)009[0699:ITII2.0.CO;2]

Mayfield, M. M., & Levine, J. M. (2010). Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology Letters*, 13(9), 1085–1093. https://doi.org/10.1111/j.1461-0248.2010.01509.x

Mazel, F., Pennell, M. W., Cadotte, M. W., Díaz, S., Dalla Riva, G. V., Grencher, R., ... Pearse, W. D. (2018). Prioritizing phylogenetic diversity captures functional diversity unreliably. *Nature Communications*, 9(1), 2888. https://doi.org/10.1038/s41467-018-05126-3

Mc Cannon, K. S. (2000). The diversity–stability debate. *Nature*, 405(May), 228–233. https://doi.org/10.1038/35012234

McDonald, D., Clemente, J. C., Kuczynski, J., Rideout, J. R., Stombaugh, J., Wendel, D., ... Caporaso, J. G. (2012). The Biological Observation Matrix (BIOM) format or: How I learned to stop worrying and love the ome-ome. *GigaScience*, 1(1), 464. https://doi.org/10.1186/2047-217X-1-7

McFarlin, K. M., Questel, J. M., Hopcroft, R. R., & Leigh, M. B. (2017). Bacterial community structure and functional potential in the northeastern Chukchi Sea. *Continental Shelf Research*, 136, 20–28. https://doi.org/10.1016/j.csr.2017.01.018

McGee, K. M., Robinson, C. V., & Hajibabaei, M. (2019). Gaps in DNA-based biomonitoring across the globe. *Frontiers in Ecology and Evolution*, 7, 337. https://doi.org/10.3389/frse.2019.00337

Momal, R., Robin, S., & Ambroise, C. (2019). Tree-based reconstruction of ecological network from abundance data. *ArXiv*, 1–23. Retrieved from http://arxiv.org/abs/1905.02452

Morris, Z. S., Wooding, S., & Grant, J. (2011). The answer is 17 years, what is the question: Understanding time lags in translational research. *Journal of the Royal Society of Medicine*, 104(12), 510–520. https://doi.org/10.1258/jrsm.2011.110180

Mortágua, A., Vasselon, V., Oliveira, R., Elias, C., Chardon, C., Bouchez, A., ... F.P. Almeida, S. (2019). Applicability of DNA metabarcoding approaches including predictive metagenomic profiling reveal characteristics of bacterial response to petroleum hydrocarbon contamination in diverse environments. *Scientific Reports*, 7(1), 1108. https://doi.org/10.1038/s41598-017-01126-3

Nascimento, F. A. J., Lallias, D., Blik, H. M., & Creer, S. (2018). Sample size effects on the assessment of eukaryotic diversity and community structure in aquatic sediments using high-throughput sequencing. *Scientific Reports*, 8(1), 11737. https://doi.org/10.1038/s41598-018-30179-1

Nguyen, T. T., Cochrane, S. K. J., & Landfald, B. (2018). Perturbation of seafloor bacterial community structure by drilling waste discharge. *Marine Pollution Bulletin*, 129(2), 615–622. https://doi.org/10.1016/j.marpolbul.2017.10.039

Niemeyer, D. (2002). Developing indicators for environmental policy: Data-driven and theory-driven approaches examined by example. *Environmental Science and Policy*, 5(2), 91–103. https://doi.org/10.1016/S1462-9011(02)00026-6

Nilsson, L. K. J., Sharma, A., Bhatnagar, R. K., Bertilsson, S., & Tenerius, O. (2019). Presence of *Aedes* and *Anopheles* mosquito larvae is correlated to bacteria found in domestic water-storage containers. *FEMS Microbiology Ecology*, 94(6), fiy058. https://doi.org/10.1093/femsec/fiy058

Nilsson, R. H., Larsson, K.-H., Taylor, A. F. S., Bengtsson-Palme, J., Jeppesen, T. S., Schigel, D., ... Abarenkov, K. (2019). The UNITE database for molecular identification of fungi: Handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Research*, 47(D1), D259–D264. https://doi.org/10.1093/nar/gky1022

Obi, C. C., Adebuseoye, S. A., Ugji, E. O., Ilori, M. O., Ampund, O. O., & Hickey, W. J. (2016). Microbial communities in sediments of Lagos Lagoon, Nigeria: Elucidation of community structure and potential impacts of contamination by municipal and industrial wastes. *Frontiers in Microbiology*, 7(AUG), 1213. https://doi.org/10.3389/fmicb.2016.01213

Oliver, T. H., Heard, M. S., Isaac, N. J. B., Roy, D. B., Procter, D., Eigenbrod, F., ... Bullock, J. M. (2015). Biodiversity and resilience of ecosystem functions. *Trends in Ecology and Evolution*, 30(11), 673–684. https://doi.org/10.1016/j.tree.2015.08.009

Pagenkopp Lohan, K. M., Fleischer, R. C., Carney, K. J., Holzer, K. K., & Ruiz, G. M. (2016). Amplicon-based pyrosequencing reveals high diversity of protistan parasites in ships’ ballast water: Implications for biogeography and infectious diseases. *Microbial Ecology*, 71(3), 530–542. https://doi.org/10.1007/s00248-015-0684-6

Pauvert, C., Vallance, J., Delière, L., Buée, M., & Vacher, C. (2019). Microbial networks inferred from metabarcoding data lack replicability: Consequences for next-generation biomonitoring. *BioRxiv*, 642199.https://doi.org/10.1101/642199

Pawlowski, J., Esling, P., Lejzerowicz, F., Cederhagen, T., & Wilding, T. A. (2014). Environmental monitoring through protist next-generation sequencing metabarcoding: Assessing the impact of fish farming on benthic foraminifera communities. *Molecular Ecology Resources*, 14(6), 1129–1140. https://doi.org/10.1111/1755-0998.12261

Pawlowski, J., Esling, P., Lejzerowicz, F., Corderie, T., Visco, J. A., Martins, C., ... Cederhagen, T. (2016). Benthic monitoring of salmon farms in Norway using foraminiferal metabarcoding. *Aquaculture Environment Interactions*, 8, 371–386. https://doi.org/10.3354/aei00182

Pawlowski, J., Kelly-Quinn, M., Altermatt, F., Apotheloz-Perret-Gentil, L., Beja, P., Boggero, A., ... Kahler, M. (2018). The future of biotic indices in the ecogenomic era: Integrating (e)DNA metabarcoding in biological assessment of aquatic ecosystems. *Science of the Total Environment*, 637–638, 1295–1310. https://doi.org/10.1016/j.scitenv.2018.05.002

Payne, R. J. (2013). Seven reasons why protists make useful bioindicators. *Acta Protozoologica*, 52(3), 105–113. https://doi.org/10.4467/1689027AP.13.0011.1108
Vasselon, V., Domaizon, I., Rimet, F., Kahler, M., & Bouchez, A. (2017). Application of high-throughput sequencing (HTS) metabarcoding to diatom biomonitoring: Do DNA extraction methods matter? *Freshwater Science*, 36(1), 162–177. https://doi.org/10.1086/690649

Vasselon, V., Rimet, F., Tapolczai, K., & Bouchez, A. (2017). Assessing ecological status with diatoms DNA metabarcoding: Scaling-up on a WFD monitoring network (Mayotte Island, France). *Ecological Indicators*, 82, 1–12. https://doi.org/10.1016/j.ecolind.2017.06.024

Venail, P. A., & Vives, M. J. (2013). Phylogenetic distance and species richness interactively affect the productivity of bacterial communities. *Ecology*, 94(11), 2529–2536. https://doi.org/10.1890/12-2002.1

Větrovský, T., & Baldrian, P. (2013). The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. *PLoS One*, 8(2), e57923. https://doi.org/10.1371/journal.pone.0057923

Visco, J. A., Apothéloz-Perret-Gentil, L., Cordonier, A., Epling, P., Pillet, L., & Pawlowski, J. (2015). Environmental Monitoring: Inferring the Diatom Index from Next-Generation Sequencing Data. *Environmental Science and Technology*, 49(13), 7597–7605. https://doi.org/10.1021/es506158m

Vivien, R., Lejezerowicz, F., & Pawlowski, J. (2016). Next-generation sequencing of aquatic oligochaetes: Comparison of experimental communities. *PLoS One*, 11(2), e0148644. https://doi.org/10.1371/journal.pone.0148644

von Ammon, U., Wood, S. A., Larche, O., Zaiko, A., Lavery, S. D., Inglis, G. J., & Pochon, X. (2019). Linking environmental DNA and RNA for improved detection of the marine invasive fanworm Sabella spallanzanii. *Frontiers in Marine Science*, 6, 621. https://doi.org/10.3389/fmars.2019.00621

Washburne, A. D., Morton, J. T., Sanders, J., McDonald, D., Zhu, Q., Oliverio, A. M., & Knight, R. (2018). Methods for phylogenetic analysis of microbiome data. *Nature Microbiology*, 3(6), 652–661. https://doi.org/10.1038/s41564-018-0156-0

Waters, C. N., Zalasiewicz, J., Summerhayes, C., Barnosky, A. D., Poirier, C., Ga uszka, A., ... Wolfe, A. P. (2016). The Anthropocene is functionally and stratigraphically distinct from the Holocene. *Science*, 351(6269), aad2622. https://doi.org/10.1126/science.aad2622

Webb, C. O., Ackerly, D. D., McPeek, M. A., & Donoghue, M. J. (2002). Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, 33(2002), 475–505. https://doi.org/10.1146/annurev.ecolsys.33.010802.150448

Weigand, H., Beermann, A. J., Ciampor, F., Costa, F. O., Csabai, Z., Duarte, S., ... Ekrem, T. (2019). DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. *Science of the Total Environment*, 678, 499–524. https://doi.org/10.1016/j.scitotenv.2019.04.247

Welsh, H. H., & Ollivier, L. M. (1998). Stream amphipods as indicators of ecosystem stress: A case study from California’s redwoods. *Ecological Applications*, 8(4), 1118–1132.

Wilcox, T. M., Carim, K. J., Young, M. K., McKelvey, K. S., Franklin, T. W., & Schwartz, M. K. (2018). Comment: The importance of sound methodology in environmental DNA sampling. *North American Journal of Fisheries Management*, 38(3), 592–596. https://doi.org/10.1002/nafm.10055

Xu, Z., Malmer, D., Langille, M. G. I., Way, S. F., & Knight, R. (2014). Which is more important for classifying microbial communities: Who’s there or what they can do? *ISME Journal*, 8(12), 2357–2359. https://doi.org/10.1038/ismej.2014.157

Yilmaz, P., Kottmann, R., Field, D., Knight, R., Cole, J. R., Amaral-Zettler, L., ... Glöckner, F. O. (2011). Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIXS) specifications. *Nature Biotechnology*, 29(5), 415–420. https://doi.org/10.1038/nbt.1823

Zaiko, A., Samuiloviene, A., Ardua, A., & Garcia-Vazquez, E. (2015). Metabarcoding approach for nonindigenous species surveillance in marine coastal waters. *Marine Pollution Bulletin*, 100(1), 53–59. https://doi.org/10.1016/j.marpolbul.2015.09.030

Zappelini, C., Karimi, B., Foulon, J., Lacerca-Díaz, L., Maillard, F., Valot, B., ... Chalot, M. (2015). Diversity and complexity of microbial communities from a chlor-alkali tailings dump. *Soil Biology and Biochemistry*, 90, 101–110. https://doi.org/10.1016/j.soilbio.2015.08.008

Zhou, J., Deng, Y., Luo, F., Molecular, P., Network, E., Communities, S. M., ... Journal, A. S. M. (2011). Phylogenetic molecular ecological network of soil microbial. *MBio*, 2(4), 1–8. https://doi.org/10.1128/mBio.00122-11. Editor

Zhou, X., Li, Y., Liu, S., Yang, Q., Su, X. U., Zhou, L., ... Huang, Q. (2013). Ultra-deep sequencing enables high-fidelity recovery of biodiversity for bulk arthropod samples without PCR amplification. *GigaScience*, 2(1), 4. https://doi.org/10.1186/2047-217X-2-4

Zimmermann, J., Abarca, N., Enke, N., Enk, N., Skibbe, O., Kusber, W. H., & Jahn, R. (2014). Taxonomic reference libraries for environmental barcoding: A best practice example from diatom research. *PLoS One*, 9(9), e108793. https://doi.org/10.1371/journal.pone.0108793

Zinger, L., Bonin, A., Alsos, I. G., Bálint, M., Bik, H., Boyer, F., ... Taberlet, P. (2019). DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions. *Molecular Ecology*, 28(8), 1857–1862. https://doi.org/10.1111/mec.15060

**SUPPORTING INFORMATION**

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

[How to cite this article: Cordier T, Alonso-Sáez L, Apothéloz-Perret-Gentil L, et al. Ecosystems monitoring powered by environmental genomics: A review of current strategies with an implementation roadmap. *Mol Ecol*. 2021;30:2937–2958.](https://doi.org/10.1111/mec.15472)