Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

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

Tristan Cordier[1*], Laura Alonso-Sáez[2], Laure Apothéloz-Perret-Gentil[1], Eva Aylagas[3], David A. Bohan[4], Agnès Bouchez[5], Anthony Chariton[6], Simon Creer[7], Larissa Frühe[8], François Keck[5], Nigel Keeley[10], Olivier Laroche[10], Florian Leese[11,12], Xavier Pochon[13,14], Thorsten Stoeck[8], Jan Pawlowski[1], Anders Lanzén[2,15]

[1]Department of Genetics and Evolution, University of Geneva, Science III, 4 Boulevard d'Yvoy, 1205 Geneva, Switzerland

[2]AZTI, Txatxarramendi Irla s/n, 48395 Sukarrieta, Spain

[3] Red Sea Research Center (RSRC), Biological and Environmental Sciences and Engineering (BESE), King Abdullah University of Science and Technology (KAUST), Saudi Arabia

[4]Agroécologie, AgroSup Dijon, INRAE, Univ. Bourgogne, Univ. Bourgogne Franche-Comté, 17 rue Sully, BP 86510, F-21000, Dijon, France

[5]UMR CARRTEL, INRAE, F-74200 Thonon, France

[6]Department of Biological Sciences, Macquarie University, NSW, Australia 2113

[7]School of Natural Sciences, Bangor University, Gwynedd, LL57 2UW, UK

[8]Department of Ecology, Technische Universität Kaiserslautern, Erwin-Schroedinger-Str. 14, D-67663 Kaiserslautern, Germany

[10]Benthic Resources and Processes Group, Institute of Marine Research, Tromsø, Norway

[11]Aquatic Ecosystem Research, Faculty of Biology, University of Duisburg-Essen, Universitaetsstr. 5, 45141 Essen, Germany

[12]Centre for Water and Environmental Research (ZWU), University of Duisburg-Essen, Universitaetsstr. 2, 45141 Essen, Germany

[13]Coastal & Freshwater Group, Cawthron Institute, 98 Halifax Street East, 7010 Nelson, New Zealand

[14]Institute of Marine Science, University of Auckland, Warkworth 0941, New Zealand

[15]IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

Contact: tristan.cordier@gmail.com

Key words: Biomonitoring; metabarcoding; environmental DNA; biodiversity; implementation strategy; ecosystem management

Abstract

A decade after environmental scientists integrated high-throughput sequencing technologies in their toolbox, the genomics-based monitoring of anthropogenic impacts on biodiversity and ecosystems is yet to be implemented by regulatory frameworks. Despite the broadly acknowledged potential of environmental DNA and RNA to cost-efficiently and accurately monitor biodiversity, technical limitations and conceptual issues still stand in the way of its routine application by end-users. In addition, the multiplicity of potential implementation strategies may contribute to a perception of the methodology as being premature or "in development", hence restraining regulators from binding these tools into legal frameworks. This review focuses on the strengths and limitations of genomics-based strategies that have emerged over the past ten years and have been classified for this purpose into three broad strategies: (A) Taxonomy-based approaches that focus on known bio-indicators or the diversity of taxonomically described taxa, (B) *De novo* approaches that do not require well-established taxonomy, and (C) Function-based approaches that rely on community-wide metrics, where taxa are interchangeable, or on functional profiles instead of compositional turnovers. We finally propose a roadmap for the implementation of environmental genomics into routine monitoring programs that leverage recent analytical advancements, upon which some critical limitations are alleviated.

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

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

of biodiversity across the globe (Butchart et al., 2010; Cardinale et al., 2012; Hughes et al. 2018). Recent reports stress that one million species out of the 8 million known are presently at risk of extinction (IPBES report, 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 in real time in order to forecast the direction of those shifts and their impacts on ecosystem functions and services (Carpenter et al., 2011; Barnosky et al., 2012; Ratajczak et al., 2018).

Because our societies are aiming to reach a trade-off between socioeconomic 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; Grizetti et al., 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, UNCLOS). The backbone of such monitoring programs is the biological component of ecosystems, as a measure of ecosystem 'health' or 'integrity' (Karr, 1999), and referred to as the Biological Quality Elements (BQEs, Borja et al., 2013; Hering et al., 2018). Most monitoring strategies implemented in regulations rely on the bioindication principle (autecology), that is, a significant correlation between the occurrence of specific organisms and a set of environmental variables. Indeed, 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 (Carignan & Villard, 2002; Lear et al., 2011; Birk et al., 2012).

The limits of currently implemented ecosystem monitoring strategies

Traditionally, morphologically distinguishable invertebrates have been used as bioindicators in both aquatic and terrestrial ecosystems (Reynoldson & Metcalfe-Smith, 1992; Bongers & Ferris, 1999; Hodkinson & Jackson, 2005; Gerlach et al., 2013) while fishes, amphibians, macrophytes, phytoplankton and diatoms are also routinely used in aquatic

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 January 2020 doi:10.20944/preprints202001.0278.v1

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

ecosystems (Birk et al., 2012). Various Biotic Indices (BIs) have been formalized, based on the predictable responses of species to environmental disturbances (autecological value) in marine (Maurer et al., 1999; Borja et al., 2000; Rygg et al., 2013), aquatic inland (Kelly et al., 1995; Stark et al., 1998; Prygiel & Coste, 2000) and terrestrial ecosystems (Urzelai et al., 2000; Marull et al., 2007). Almost half of the monitoring methodologies currently used in Europe are underpinned by such BIs (Birk et al., 2012). However, for environments or geographical regions for which no BI has been calibrated, ecological assessments rely instead on biodiversity measures of "charismatic" groups such as fishes (Pont et al., 2006), amphibians (Welsh et al., 1998) and insects (Basset el al., 2004).

These monitoring methodologies require the collection and identification of hundreds to thousands of specimens per sample, which is a slow, labor-intensive process. These limitations seriously hamper our capacity to scale up biomonitoring and satisfy the increasing demand for environmental monitoring programs in a timely fashion that allows informed ecosystem management (Baird & Hajibabaei, 2012). Moreover, this conventional morphology-based approach is compromised by several other shortcomings, (i) it focuses only on morphologically identifiable biodiversity, ignoring inconspicuous meiofaunal and microbial taxa, which are known to be strong bioindicators; (ii) cryptic diversity (morphological look-alikes with a variable tolerance to disturbances) remains unrecognized; (iii) variation in species life stages, damaged specimens and misidentifications caused by decreasing taxonomic expertise worldwide may lead to variable and noisy species' inventories, and by extension, to ecological assessments. Taken together, the need for faster, more objective, robust and cost-effective tools and strategies to deliver a more efficient ecosystem monitoring has never been more pressing.

The 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; Valentini et al., 2016; Deiner et al., 2017; Cristescu et al., 2018; Ruppert et al., 2019) and their metabolic functions (metagenomics and metatranscriptomics, reviewed in Ungerer et al., 2008; Vandenkoornhuyse et al., 2010; Quince

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

et al., 2017; Singer et al., 2017; Escalas 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 (Ratnasingham & 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 unraveled 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 (Delgado-Baquerizo et al., 2016; Guidi et al., 2016; Cavicchioli et al., 2019), the ecology of most microorganisms remains largely enigmatic.

The potential of EG for surveying biodiversity and monitoring natural ecosystems at a broad spatio-temporal scale was quickly identified and implemented by environmental scientists (Baird & Hajibabaei, 2012; Taberlet et al., 2012; Davies et al., 2012; Kelly et al., 2014). This work was boosted by the massive drop in cost of sequencing technologies over four orders of magnitude within the last 15 years (https://www.genome.gov), promoting numerous clinical and environmental routine applications. Indeed, fueled by the continuous efforts to optimize laboratory protocols and bioinformatic tools, the large-scale collection of samples, the generation of HTS data, the statistical analysis of the data and the interpretation of the results can now be performed in near real time (Juul et al., 2015; Quinn et al., 2016; Deshpande et al., 2019). The next breakthrough of the EG 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) that may be integrated to autonomous instruments for broad-scale and continuous ecosystem monitoring programs (Brandt et al., 2016; Bohan et al., 2017; Aguzzi et al., 2019; Benway et al., 2019; Levin et al., 2019).

These advances in genomics-based research led to a series of studies to assess the applicability of EG for the monitoring of ecosystem changes by collecting biodiversity data on various biological models (fishes, macroinvertebrates, protists, bacteria) populating various environments (water, biofilms, sediment). While some of these pilot studies targeted multicellular organisms in replacement for arduous morphological identification (Hajibabaei et

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 January 2020 doi:10.20944/preprints202001.0278.v1

Peer-reviewed version available at Molecular Ecology 2020; doi:10.1111/mec.15472

al., 2011, 2012; Thomsen et al., 2012; Zhou et al., 2013; Lejzerowicz et al., 2015), the potential of EG to leverage the general eukaryotic and prokaryotic diversity for ecological monitoring was also explored (Chariton et al., 2010; Bik et al., 2012; Dowle et al., 2015; Lallias et al., 2015) and advocated (Creer et al., 2010; Payne, 2013; Bouchez et al., 2016; Chariton et al., 2016; Graham et al., 2016). Given the immense opportunities opened by EG for ecosystem monitoring, over 45 countries recently decided to join their efforts within the DNAqua-Net European COST Action, 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 projects were recently launched, such as STREAM in Canada (https://stream-dna.com/), Lakes380 in New Zealand (https://lakes380.com/) and NGB in France (http://next-genbiomonitoring.org/), all aiming at the unbridling of EG for ecosystem monitoring. In fact, multiple pilot and methodological EG studies have highlighted important variation in terms of EG applicability to comply with the criteria laid down in current regulatory programs (reviewed in Hering et al., 2018), leading to the proposition of multiple implementation scenarios for current and future ecosystem monitoring programs.

In this paper, we review the strengths and limitations of the different EG implementation strategies and propose to classify them into three categories (Figure 1, Table 1), (A) Taxonomy-based approaches that focus on known bio-indicators or the diversity of formally or informally described taxa, (B) *De novo* approaches aiming to identify and utilise novel bioindicators, often independent from formal taxonomy, and (C) Function-based approaches that rely on community-wide metrics or functional profiles instead of taxonomic composition. Based on the specificities of each strategy, their level of maturity and their compatibility with existing regulations (see Table 2), we propose an implementation roadmap relying on recent analytical advancements that alleviate some critical limitations.

Strategy A ("Taxonomy-based"): Screening known species and bioindicators with environmental genomics

This strategy relies on two basic endeavours: either the detection of DNA traces that morphologically distinguishable species leave behind in an environmental sample (sediment or water), or DNA obtained from bulk material prepared from an environmental sample by e.g. elutriation, trapped individuals or biofilm scratching (Figure 1A). This strategy closely fits the

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

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, since these are usually unidentifiable on the basis of morphological traits, and for most of which the autoecology is largely unknown (diatoms being an exception). The reference databases routinely used by EG studies include for instance the universal but essentially non-curated GenBank nucleotide repository from the National Center for Biotechnology Information (Benson et al., 1999, but see Leray et al., 2019), or the curated databases BOLD for COI barcodes primarily from animals (Ratnasingham & Hebert, 2007), SILVA for universal ribosomal markers (Quast et al., 2013), PR^2 for protists (Guillou et al., 2013), Diat.barcode for diatoms (Rimet et al., 2016), and Unite for fungi (Nilsson et al., 2018).

Depending on the environmental matrix assessed and the taxonomic group considered, the performance of EG-based methodologies for monitoring varies considerably (Hering et al., 2018). Benchmarking studies comparing EG-based taxonomic inventories and conventional morphology-based surveys have shown mixed degrees of congruence. For the non-invasive detection of species DNA traces in filtered marine water, the rate of success from EG-based monitoring was 100% (Thomsen et al., 2012; Bakker et al., 2017). For freshwater macroinvertebrate bulk samples, the rate of species detection varied from 67% (Elbrecht et al., 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., 2017; Vasselon et al., 2017a) to 28% (Visco et al., 2015). The previously reported congruence for macroinvertebrates sampled from marine sediments ranged from 20% (Lejzerowicz et al., 2015) up to 60% (Aylagas et al., 2016). Noteworthy, those studies also detected numerous species that were unnoticed in morphological inventories (Hajibabaei et al., 2011; 2012; Elbrecht et al., 2017). Despite these discrepancies, the studies inferring BI values from the detected bioindicators species showed very promising results, for both freshwater diatoms (Kermarrec et al., 2014; Visco et al., 2015; Vasselon et al., 2017b; Kelly et al., 2018) and macroinvertebrates (Elbrecht et al., 2017) as well as for marine macroinvertebrates (Lejzerowicz et al., 2015; Aylagas et al., 2016). While

Peer-reviewed version available at Molecular Ecology 2020; doi:10.1111/mec.15472

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 (Aylagas 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. Likewise, further improvements in molecular protocols as well as BI inter-calibration is a necessity towards harmonization and standardization across Europe (Poikane et al., 2014; Hering et al., 2018) 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). The limitations mainly stem from the fact that the methods are sampling fundamentally different units (molecules versus 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 et al., 2017), should not be considered analogous to morpho-species richness even in the absence of noise that can result from PCR and sequencing. This discrepancy is due to cryptic diversity (Stork, 2018), intragenomic marker variation (Bik et al., 2013, Sun et al., 2013) and the presence of DNA from dead and inactive organisms or as extracellular molecules (Collins et al., 2018). The abundance of taxa inferred from HTS read counts should also not be used to derive the number of individuals. Indeed, the number of sampled DNA molecules and HTS read counts 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, Vetrovský, et al., 2013), in addition to variations in DNA extractability and primer-specific amplification bias (Elbrecht et al., 2015; Piñol et al., 2015; Krehenwinkel et al., 2017). 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 et al., 2017). However, bulk samples (Elbrecht et al., 2017), larger extraction volume (Nascimento et al., 2018) or more aggressive homogenization (Lanzén et al., unpublished data) can partially alleviate this issue.

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

As taxonomy-based strategy (A) depends on reference sequences for organism identification, the incompleteness of reference databases can 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 (Vasselon et al., 2017; McGee et al., 2019). However, despite sustained efforts, reference databases will likely remain skewed towards some taxa while suffering from important gaps across other taxonomic groups and/or biogeographical regions (Weigand et al., 2019). All those issues directly impact both qualitative and quantitative measures of biodiversity using taxonomy-based EG approaches, two key parameters for the calculation of BIs (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 (Vetrovský, et al., 2013) or by spiking samples with known internal standard for quantitative determinations (Tkacz et al., 2018; Ji et al., 2019), are all promising methods for resolving these challenges. Finally, the integration of bioinformatic tools for the automatized curation of databases from mislabeled sequences will improve their reliability (Ashelford et al., 2005; Kozlov et al., 2016).

Strategy B ("*de novo*"): Discovering new bioindicators and harnessing them for routine monitoring.

In contrast to the taxonomy-based Strategy A, Strategy B does not immediately generate an ecological assessment, because it does not use previous knowledge of morphology-based bioindicators. Instead, this *de novo* strategy aims at establishing new bioindicators using EG-based profiling of community structures and independently generated ecological status or known disturbance gradients (Figure 1B). Harnessing EG with HTS technologies to explore a broader range of biological diversity, formally labelled or not (i.e. taxonomically assigned or not), represents an opportunity to move towards a more holistic

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 January 2020 doi:10.20944/preprints202001.0278.v1
Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

monitoring methodology (Chariton et al., 2010; Bik et al., 2012). By considering all the OTUs (or ASVs) profiles along a known impact gradient of typical anthropogenic origin, these studies have shown that HTS data represent a virtually unlimited reservoir of new bioindicators. Examples (listed in Table 1) include contamination by pesticide (Thompson et al., 2016; Andùjar et al., 2017), agriculture stressors (Salis et al., 2017) and gradients of eutrophication and urban contamination (Apothéloz et al., 2017; Martínez-Santos et al. 2018; Simonin et al., 2019; Tapolczai et al., 2019a, 2019b) in freshwater systems. In marine environments, it was demonstrated after an oil spill (Bik et al., 2012), in the vicinity of offshore drilling platforms (Lanzén et al., 2016; Laroche et al., 2016, 2018a) and aquaculture sites (Pochon et al., 2015; Dowle et al., 2015; Keeley et al., 2018; Stoeck et al. 2018a, 2018b) as well as along eutrophication and urban or industrial contamination gradients in estuaries (Chariton et al., 2010, 2015; Angly et al., 2015; 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 (Chariton et al., 2010; Bik et al., 2012; Lanzén et al., 2016), large groups of protists such as diatoms and ciliates (Lanzén et al., 2016; Stoeck et al., 2018a) or foraminifera (Pawlowski et al., 2014; Laroche et al., 2016; Frontalini et al., 2018) as well as microbial groups such as fungi (Bik et al., 2012), oomycetes (Lanzén et al., 2016) and bacteria (Angly et al., 2015; Dowle et al., 2015; Martínez-Santos et al. 2018; Obi et al., 2016; Aylagas et al., 2017; Stoeck et al., 2018b; Keeley et al. 2018) have great potential as bioindicators of anthropogenic impacts and that can readily be captured by EG studies.

Unfortunately, most of these proof-of-concept studies have not yet validated their results either by performing ecological assessments based on identified indicators as a reference in new environments or existing public sequence datasets from external sites. For this information to be useful in ecosystem monitoring on new samples, the data obtained from along known disturbance gradients (i.e. reference or training dataset) must be operational in different spatiotemporal contexts. To this end, two main approaches have been proposed and tested, namely the indicator value (e.g. the IndVal approach, Dufrêne and Legendre 1997) and supervised machine learning (SML, Crisci et al., 2012; Libbrecht & Noble, 2015).

The indicator value approach, which is conceptually similar to the BIs, formalized using morphology-based data, ascribes autoecological values (or discrete “eco-groups”) to OTUs or ASVs based on their occurrence in samples of known disturbance level. Hence, the resulting

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

molecular-based index is directly calibrated on the HTS data, which alleviates the qualitative and quantitative biases encountered with the taxonomy-based EG approaches. Such strategy has proven successful for both freshwater benthic diatoms (Apothéloz et al., 2017; Tapolczai et al., 2019a, 2019b) and for water bacterial and eukaryotic communities in streams and estuarine systems (Chariton et al., 2015; Li et al., 2018). A strategy analogous to the IndVal approach is the use of polynomial quantile regression splines (Andersson, 2008), which has shown great promise for the prediction of 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 et al., 2017; Tapolczai et al. 2019a, 2019b).

The supervised machine learning (SML) approach also uses training datasets, i.e. the reference disturbance level (labels) associated with the OTU (or ASV) profiles of the samples (features). Supervised machine learning algorithms are best at classification problems involving multidimensional and noisy datasets (Libbrecht et al., 2015), which are common attributes of HTS data. The task is to automatically disentangle the feature signal (OTU or ASV) and their associations that convey an ecological signal, explaining the disturbance level, 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 et al., 2019a). This methodology also alleviates the qualitative and quantitative biases that hamper taxonomy-based approaches, because the model is trained directly on the HTS data. The applicability of SML has been demonstrated in marine environments, for the detection of various pollutants, the prediction of geochemical features (Smith et al., 2015) and for the prediction of aquaculture impacts on benthic diversity (Cordier et al., 2017; 2018). The SML-based inference of BI values was also found to outperform the taxonomy-based strategy, that relies only on the detection of established macroinvertebrates bioindicators DNA (Cordier et al., 2018), and can be more powerful that the IndVal approach (Frühe et al., submit. in this issue). Supervised machine learning applications have also succeeded in predicting the origin of container ship ballast waters (Gerhard & Gunsch, 2019).

The *de novo* strategy potentially provide numerous advantages over the taxonomy-based one. First, the dependence on reference sequence databases for taxonomic assignments

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 January 2020 doi:10.20944/preprints202001.0278.v1

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

of HTS reads to known bioindicators is reduced or even bypassed. Instead, new ecological knowledge is hypothesised during the formulation or calibration of a molecular index (indicator value) or during the supervised training of a model (SML). Second, it also leverages previously inaccessible powerful bioindicators such as microorganisms, protists, meiofauna and mesozooplankton, that are widespread and may react both faster and stronger to environmental disturbances (Creer et al., 2010; Payne, 2013; Bouchez et al., 2016). Finally, when applied for the inference of BIs that are currently implemented in routine ecological monitoring programs, the *de novo* strategy that is calibrated on these BIs is directly compatible with current regulations, because the evaluation criteria, i.e. the BI, remain unchanged. Hence, this strategy assure a full backward and forward compatibility with current monitoring programs, facilitating continuity of important time series datasets (Bálint et al., 2018).

Strategy C ("function-based"): Blending theoretical and functional ecology into routine ecosystem monitoring.

Strategy C relies on metrics derived from the community structure or its functional profiles – where taxa are interchangeable – in order to infer the impact of disturbance and its ramifications on ecosystem functioning (Figure 1C). This represents the clearest paradigm-shift for ecosystem monitoring programs, because the evaluation of bioindicators, based on the compositional variation of taxa, is not the main aim of the strategy. Instead, its focus is to discover and understand the 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-independent biodiversity metrics and ecosystems functioning and resilience, to reach a more general theoretical framework (Cardinale et al., 2000; McCann, 2000; Hooper et al., 2005; Tilman et al., 2006; Ives & Carpenter, 2007; Mouillot et al., 2013; Loreau & de Mazancourt, 2013).

Community-wide metrics can be computed from compositional data generated by EG studies, including alpha-diversity (i.e. OTUs or ASVs richness, diversity or evenness; reviewed in Daly et al., 2018), and these can be considered a simplistic version of "function-based" strategy, along with their phylogeny-aware derivatives (reviewed in Tucker et al., 2017; Washburne et al., 2018). Under anthropogenic impact, alpha-diversity has been found to

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 January 2020 doi:10.20944/preprints202001.0278.v1

Peer-reviewed version available at Molecular Ecology 2020; doi:10.1111/mec.15472

decrease for the foraminifera (Pawlowski et al., 2014; 2016; Laroche et al. 2018b), ciliates (Stoeck et al., 2018a) and bacterial communities (Stoeck et al., 2018b) of marine sediments. 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 an 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 et al., 2002; Cavender-Bares et al., 2009, but see Mayfield & Levine, 2010; Gerhold et al., 2015). This relationship between phylogenetic diversity and ecosystem functioning has received a lot of attention by plant ecologists (Flynn et al., 2011). However, only few studies have employed EG data, targeting mostly microbial groups, to empirically explore this relationship, which, as for simple alpha-diversity metrics, has resulted in contrasting conclusions (Galand et al., 2015; Pérez-Valera et al., 2017, Liu et al., 2017; but see Venail & Vives, 2013; Keck & Kahlert, 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 non-linear, 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 (Webb et al., 2002; Cavender-Bares et al., 2009; Srivastava et al., 2012). 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 et al., 2012; Oliver et al., 2015). However, because not all functional traits necessarily have a phylogenetic signal (Srivastava et al., 2012), including for microbes (Martiny et al., 2013), inferring ecosystem functioning and the level of anthropogenic impact based on phylogeny-aware metrics alone may

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 January 2020 doi:10.20944/preprints202001.0278.v1

Peer-reviewed version available at Molecular Ecology 2020; doi:10.1111/mec.15472

prove to be misguided. Likewise, conservation strategies based on these metrics may also be suboptimal (Mazel et al., 2018).

Another set of community-wide metrics can be computed from the structure of biotic interactions within ecological networks (reviewed in Faust & Raes, 2012; Vacher et al., 2016; Layeghifard et al., 2017). Based on empirical evidence of the variation in network structure under environmental disturbance (Tylianakis et al., 2007; Zhou et al., 2011; Karimi et al., 2016; Ma et al., 2018), their properties have been suggested as potential indicators of ecosystem functioning and integrity (Gray et al., 2014; Karimi et al., 2017; Bohan et al., 2011, 2017; Lau et al., 2017; Tylianakis et al., 2017; Pellissier et al., 2018; Delmas et al., 2019). In recent years, a growing interest in these approaches has led to a series of studies employing EG to infer ecological networks from microbial communities data (Zhou et al., 2011; Lupatini et al., 2014; Zappelini et al., 2015; Pérez-Valera et al., 2017; Pauvert et al., 2019) and from macroinvertebrates (Compson et al., 2019) to explore the links between network properties (e.g. connectance, centrality, nestedness) and ecosystem functions. For instance, it has been shown that anthropized soil bacterial communities may have fewer potentially interacting taxa within the community than in natural soil (Lupatini et al., 2014). In aquatic ecosystems, anthropogenic impacts can be detected in co-occurrence networks by a lower connectivity (Lawes et al., 2017; Laroche et al., 2018b; Li et al., 2018) and a lower ratio of positive interactions (Laroche et al., 2018b).

While promising, exploring the links between the properties of ecological networks inferred from EG and ecosystem functions is still in its infancy (Faust et al., 2012; 2015; Lima-Mendez et al., 2015; Lawes et al., 2017; Laroche et al., 2018b; Li et al., 2018; Pauvert et al., 2019). Multiple methodological issues still limit the inference of robust networks from EG data of co-occurrences, among which is the composition-like nature of reads counts (but see Friedman & Alm, 2012; Kurtz et al., 2015), the control for covariates and background information (but see Tammadoni-Nezhad et al., 2013; Tackmann et al., 2018; Cougoul et al., 2018; Chiquet et al., 2018; Momal et al., 2019), replicability of inference (Pauvert et al., 2019) and the relative merits of statistical and logical inference (Vacher et al., 2016) remain challenging issues. Furthermore, robust networks require considerably more replicates than typically collected in EG studies, which greatly increase both time and costs. However, as more benchmark datasets containing both EG data and independently confirmed interactions between taxa become available to

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

complement simulated datasets (see Lima-Mendez et al., 2015), making robust networks inference to explore the applicability of their metrics for ecosystem monitoring will come within reach in the years to come.

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. 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 et al., 2014). Some metagenomic contigs and functional transcripts were indeed found to represent efficient indicators of anthropogenic disturbances, in terrestrial (de Menezes et al., 2012), groundwater (He et al., 2018), freshwater (Thompson et al., 2016; Cheaib et al., 2018; Falk et al., 2019) and marine environments (Kisand et al., 2012; Galand et al., 2016; Birrer et al., 2019), opening up potential avenues for future routine ecosystem monitoring. Functional and taxonomic profiles may respond differently under anthropogenic disturbance (Cheaib et al., 2018), as well as under natural environmental variation (Barberàn et al., 2012; Louca et al., 2016a; 2016b; Louca et al., 2018). This taxon-function decoupling paves the way towards a molecular trait-based ecology (Raes et al., 2011; Lajoie & Kembel 2019).

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 microbial functional redundancy may be widespread (Louca et al., 2018; Pearman et al., 2019; but see Galand et al., 2018 and see 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

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

constitute useful early warnings for a timelier ecosystem management, especially the ones detected by mean of metatranscriptomics. However, RNA molecules are reportedly less stable than DNA, which would add challenging practical constraints that could preclude their implementation in routine ecosystem monitoring programs (but see Fordyce et al., 2013; Pochon et al., 2017; Cristescu, 2019; von Ammon et al. 2019). As a possible cost-effective “shortcut”, bacterial 16S rRNA profiles can be used to predict functional community profiles, based on evolutionary models (Langille et al., 2013; Aßhauer et al., 2015). Thus, taxonomic data could be also explored for searching potential functional indicators by this approach (Mukherjee et al., 2017; Laroche et al., 2018; Cordier 2019).

A roadmap for the implementation of environmental genomics for ecosystem monitoring

The emergence of standards for EG methodologies to be applied for monitoring programs.

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 et al., 2011). The emergence of consensual 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 programs (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

Peer-reviewed version available at Molecular Ecology 2020; doi:10.1111/mec.15472

adoption of standardized methodologies to produce, store and analyze 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 (Pollock et al., 2018; Knight et al., 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 specially bound to genomics methodologies, and 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 (Leese et al., 2018; Hering et al., 2018; Working Group CEN/TC230/WG28), for the sampling of eDNA (Dickie et al., 2018; Wilcox et al., 2018; CEN 2018a), the molecular protocols (Goldberg et al., 2016; Blackman et al., 2019) as well as for bioinformatics (Roy et al., 2018; Knight et al., 2018), data interoperability (McDonald et al., 2012; Callahan et al., 2017) and reference databases (CEN, 2018b).

Matching the right strategy to the right monitoring program.

Several monitoring programs may benefit quickly and reliably from an EG implementation, while some others may require further optimization of molecular protocols and/or adjustments of monitoring programs' assessment criteria (reviewed in Hering et al., 2018). For instance, monitoring programs relying primarily on taxonomic inventories still require further development. Biological and technical biases, which might be in the future partially alleviated, still hinder the congruence between the recovered species list and their relative abundances. Furthermore, despite the sustained effort, reference barcoding databases remain skewed toward some groups and geographical locations (Weigand et al., 2019), limiting congruence between genomic and morpho-taxonomic methodologies. Hence, implementation will require improvements of molecular protocols to generate EG data that better fit the current standards, or an adaptation of the currently implemented assessment criteria (Hering et al., 2018).

Monitoring programs relying on the screening of bioindicators for the computation of BI values are proposed as being compatible with an implementation of EG (Hering et al., 2018). Indeed, this compatibility is greatly facilitated by the fact that the assessment criteria, BIs here, are not meant to strictly rely on taxonomic inventories but rather on the autecology of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 January 2020 doi:10.20944/preprints202001.0278.v1

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

bioindicators. Hence, for taxonomy-based strategy (A), the BI formulations can compensate the impact of taxonomic mismatches between morphology and EG to some extent, because multiple taxa are ascribed to identical autecological values, conveying similar ecological signal (Keck et al., 2018). The applicability of this strategy has been demonstrated in freshwater (Elbrecht et al., 2017; Vasselon et al., 2017b; Kelly et al., 2018; Mortagua et al., 2019; Rivera et al., 2020) and in marine environments (Lejzerowicz et al., 2015; Aylagas et al., 2016). 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 unlabeled sequences. The *de novo* strategy (B) uses the occurrence of previously scrutinized unlabeled taxa in samples of known BI values or other impact measures to ascribe autecological values to sequences directly, or generate a predictive model (Apothéloz et al., 2017; Cordier et al., 2017; Tapolczai et al. 2019). Hence, these approaches are less sensitive to the biological and technical issues mentioned above. From an implementation perspective, the *de novo* strategy thus likely represent the best science at hand for the monitoring programs relying on BIs, because the inferred BI values convey the same biological meaning as with current methodologies, assuring the continuity with previous data and time series (Bálint et al., 2018).

The function-based strategy (C) represents an alternative that may ultimately lead to a more generic, broadly applicable ecological assessment framework (Bohan et al., 2017; Karimi et al., 2017; Tylianakis et al., 2017; Quince et al., 2017; Singer et al., 2017; Pellissier et al., 2018; Escalas et al., 2019). These approaches 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 disturbances scenarios. However, an operational ecosystem monitoring framework remains to be built upon this partially validated theoretical ecological work (but see Ma et al., 2018). In addition, the extractions of community-wide metrics remain active fields of ecological research and the emergence of a molecular trait-based ecology using metagenomics and metatranscriptomics profiles is at its infancy (Lajoie et al., 2019). Hence, their operational implementation and legal binding into established monitoring programs is premature but their ecological benefit should be anticipated. Nevertheless, the datasets to be produced in the

 doi:10.20944/preprints202001.0278.v1

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

course of future ecosystem monitoring campaigns that will include environmental samples for EG data production will certainly greatly contribute to push forward these possibilities.

Collecting reference data and eDNA/eRNA samples in parallel.

If EG-based strategies are to complement or replace current morphology-based ones, the prerequisite is to establish whether EG can provide a similar diagnostic, to ensure a smooth implementation and compatibility with existing time series (Leese et al., 2016; Bálint et al., 2018). This inevitably implies extensive parallel sampling of currently implemented and EG methodologies for some time, to build reference datasets on which the applicability can be assessed and the calibration with previous methodology performed (Leese et al., 2016; Keeley et al. 2018). To be reliable, such reference datasets have to cover as much as possible the 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.

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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 January 2020 doi:10.20944/preprints202001.0278.v1

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

Further research needs

The potential for EG-based strategies for ecosystems monitoring is enormous, and can presently fulfil most of the requirements of current monitoring programs. Moving towards an implementation of EG is certainly a paradigm-shift, but this technological breakthrough will overcome the limitations of current methodologies and enable the required up-scaling to meet monitoring needs in a changing world. Without doubts, DNA-based monitoring will pave the way for a more cost-effective, faster, reproducible and semi-automatable monitoring framework. Yet, several obstacles need to be overcome in the near future. Regardless of the analytical 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 governmental representatives would allow monitoring programs to more easily bridge the science-policy gap.
- A collection of reference data (morphology) and eDNA (or eRNA) in parallel, at least in a subset of reference points or during a transition period, will assure backward and forward compatibility of time series datasets, regardless of the envisioned analytical strategy to be employed in future monitoring campaigns.
- The efforts to complete reference sequence databases needs to be sustained, by adding more representatives of the known biodiversity, with a wider geographical coverage.
- In the same manner, a reference database framework for *de novo* strategies needs to be established. A key requirement is the ability to reliably compare OTUs or ASVs identified in monitoring programs 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 biomarker remains to be identified.
- For the identification of novel genetic biomarkers in complex (microbial) 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 function-based strategy, that will likely contribute to the establishment of a more broadly applicable monitoring framework and less constrained by the database and geographical coverage limitations.

 doi:10.20944/preprints202001.0278.v1

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

References

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. doi: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*. doi: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. doi: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. doi:10.1111/mec.14410

Angly, F. E., Heath, C., Morgan, T. C., Tonin, H., Rich, V., Schaffelke, B., … Tyson, G. W. (2016). Marine microbial communities of the Great Barrier Reef lagoon are influenced by riverine floodwaters and seasonal weather events. *PeerJ*, *2016*(4), e1511. doi:10.7717/peerj.1511

Apothéloz-Perret-Gentil, L., Cordonier, A., Straub, F., Iseli, J., Esling, P., & Pawlowski, J. (2017). Taxonomy-free molecular diatom index for high-throughput eDNA biomonitoring. *Molecular Ecology Resources*, *17*(6), 1231–1242. doi:10.1111/1755-0998.12668

Ashelford, K. E., Chuzhanova, N. A., Fry, J. C., Jones, A. J., & Weightman, A. J. (2005). At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies. *Applied and Environmental Microbiology*, *71*(12), 7724–7736. doi:10.1128/AEM.71.12.7724-7736.2005

Aßhauer, K. P., Wemheuer, B., Daniel, R., & Meinicke, P. (2015). Tax4Fun: Predicting functional profiles from metagenomic 16S rRNA data. *Bioinformatics*, *31*(17), 2882–2884. doi:10.1093/bioinformatics/btv287

Aylagas, E., Borja, Á., Irigoien, X., & Rodríguez-Ezpeleta, N. (2016). Benchmarking DNA metabarcoding for biodiversity-based monitoring and assessment. *Frontiers in Marine Science*, *3*(JUN), 96. doi:10.3389/fmars.2016.00096

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

Aylagas, E., Borja, Á., Muxika, I., & Rodríguez-Ezpeleta, N. (2018). Adapting metabarcoding-based benthic biomonitoring into routine marine ecological status assessment networks. *Ecological Indicators*, *95*, 194–202. doi:10.1016/j.ecolind.2018.07.044

Aylagas, E., Borja, Á., Tangherlini, M., Dell'Anno, A., Corinaldesi, C., Michell, C. T., … Rodríguez-Ezpeleta, N. (2017). A bacterial community-based index to assess the ecological status of estuarine and coastal environments. *Marine Pollution Bulletin*, *114*(2), 679–688. doi:10.1016/j.marpolbul.2016.10.050

Bagley, M., Pilgrim, E., Knapp, M., Yoder, C., Santo Domingo, J., & Banerji, A. (2019). High-throughput environmental DNA analysis informs a biological assessment of an urban stream. *Ecological Indicators*, *104*, 378–389. doi:10.1016/j.ecolind.2019.04.088

Baird, D. J., & Hajibabaei, M. (2012). Biomonitoring 2.0: A new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Molecular Ecology*, *21*(8), 2039–2044. doi:10.1111/j.1365-294X.2012.05519.x

Bakker, J., Wangensteen, O. S., Chapman, D. D., Boussarie, G., Buddo, D., Guttridge, T. L., … Mariani, S. (2017). Environmental DNA reveals tropical shark diversity in contrasting levels of anthropogenic impact. *Scientific Reports*, *7*(1), 16886. doi:10.1038/s41598-017-17150-2

Baldwin, D. S., Colloff, M. J., Rees, G. N., Chariton, A. A., Watson, G. O., Court, L. N., … Hardy, C. M. (2013). Impacts of inundation and drought on eukaryote biodiversity in semi-arid floodplain soils. *Molecular Ecology*, *22*(6), 1746–1758. doi:10.1111/mec.12190

Bálint, M., Pfenninger, M., Grossart, H. P., Taberlet, P., Vellend, M., Leibold, M. A., … Bowler, D. (2018). Environmental DNA time series in ecology. *Trends in Ecology and Evolution*. doi:10.1016/j.tree.2018.09.003

Banerji, A., Bagley, M., Elk, M., Pilgrim, E., Martinson, J., & Santo Domingo, J. (2018). Spatial and temporal dynamics of a freshwater eukaryotic plankton community revealed via 18S rRNA gene metabarcoding. *Hydrobiologia*, *818*(1), 71–86. doi:10.1007/s10750-018-3593-0

Barberàn, A., Fernàndez-Guerra, A., Bohannan, B. J. M., & Casamayor, E. O. (2012). Exploration of community traits as ecological markers in microbial metagenomes. *Molecular Ecology*, *21*(8), 1909–1917. doi:10.1111/j.1365-294X.2011.05383.x

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

Barnosky, A. D., Hadly, E. A., Bascompte, J., Berlow, E. L., Brown, J. H., Fortelius, M., … Smith, A. B. (2012). Approaching a state shift in Earth's biosphere. *Nature*, *486*(7401), 52–58. doi:10.1038/nature11018

Basset, Y., Mavoungou, J. F., Mikissa, J. B., Missa, O., Miller, S. E., Kitching, R. L., & Alonso, A. (2004). Discriminatory power of different arthropod data sets for the biological monitoring of anthropogenic disturbance in tropical forests. *Biodiversity and Conservation*, *13*(4), 709–732.

Beazley, M. J., Martinez, R. J., Rajan, S., Powell, J., Piceno, Y. M., Tom, L. M., … Sobecky, P. A. (2012). Microbial community analysis of a coastal salt marsh affected by the Deepwater Horizon oil spill. *PLoS ONE*, *7*(7), e41305. doi:10.1371/journal.pone.0041305

Beng, K. C., Tomlinson, K. W., Shen, X. H., Surget-Groba, Y., Hughes, A. C., Corlett, R. T., & Slik, J. W. F. (2016). The utility of DNA metabarcoding for studying the response of arthropod diversity and composition to land-use change in the tropics. *Scientific Reports*, *6*, 24965. doi:10.1038/srep24965

Benson, D. A., Boguski, M. S., Lipman, D. J., Ostell, J., Ouellette, B. F. F., Rapp, B. A., & Wheeler, D. L. (1999). GenBank. *Nucleic Acids Research*. doi:10.1093/nar/27.1.12

Benway, H. M., Lorenzoni, L., White, A. E., Fiedler, B., Levine, N. M., Nicholson, D. P., … Letelier, R. M. (2019). Ocean time series observations of changing marine ecosystems: An era of integration, synthesis, and societal applications. *Frontiers in Marine Science*, *6*(JUL). doi:10.3389/fmars.2019.00393

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., … Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution*, *22*(3), 148–155. doi:10.1016/j.tree.2006.11.004

Bik, H. M., Fournier, D., Sung, W., Bergeron, R. D., & Thomas, W. K. (2013). Intra-genomic variation in the ribosomal repeats of nematodes. *PLoS ONE*, *8*(10), e78230. doi:10.1371/journal.pone.0078230

Bik, H. M., Halanych, K. M., Sharma, J., & Thomas, W. K. (2012). Dramatic shifts in benthic microbial eukaryote communities following the deepwater horizon oil spill. *PLoS ONE*, *7*(6), e38550. doi:10.1371/journal.pone.0038550

Birk, S., Bonne, W., Borja, A., Brucet, S., Courrat, A., Poikane, S., … Hering, D. (2012). Three hundred ways to assess Europe's surface waters: An almost complete overview of biological methods to implement the Water Framework Directive. *Ecological Indicators*, *18*, 31–41. doi:10.1016/j.ecolind.2011.10.009

Birrer, S. C., Dafforn, K. A., Sun, M. Y., Williams, R. B. H., Potts, J., Scanes, P., … Johnston, E. L. (2019). Using meta-omics of contaminated sediments to monitor changes in pathways relevant to climate regulation. *Environmental Microbiology*, *21*(1), 389–401. doi:10.1111/1462-2920.14470

Blackman, R., Mächler, E., Altermatt, F., Arnold, A., Beja, P., Boets, P., … Deiner, K. (2019). Advancing the use of molecular methods for routine freshwater macroinvertebrate biomonitoring – the need for calibration experiments. *Metabarcoding and Metagenomics*, *3*, 49–57. doi:10.3897/mbmg.3.34735

Bohan, D. A., Caron-Lormier, G., Muggleton, S., Raybould, A., & Tamaddoni-Nezhad, A. (2011). Automated discovery of food webs from ecological data using logic-based machine learning. *PLoS ONE*, *6*(12), e29028. doi:10.1371/journal.pone.0029028

Bohan, D. A., Vacher, C., Tamaddoni-Nezhad, A., Raybould, A., Dumbrell, A. J., & Woodward, G. (2017). Next-generation global biomonitoring: large-scale, automated reconstruction of ecological networks. *Trends in Ecology and Evolution*, *32*(7), 477–487. doi:10.1016/j.tree.2017.03.001

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., … de Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology and Evolution*. doi:10.1016/j.tree.2014.04.003

Bongers, T., & Ferris, H. (1999). Nematode community structure as a bioindicator in environmental monitoring. *Trends in Ecology and Evolution*, *14*(6), 224–228. doi:10.1016/S0169-5347(98)01583-3

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. doi:10.1016/S0025-326X(00)00061-8

Borja, Angel, Elliott, M., Andersen, J. H., Cardoso, A. C., Carstensen, J., Ferreira, J. G., … Zampoukas, N. (2013). Good environmental status of marine ecosystems: What is it and how do we know when we have attained it? *Marine Pollution Bulletin*, *76*(1–2), 16–27. doi:10.1016/j.marpolbul.2013.08.042

Borrell, Y. J., Miralles, L., Do Huu, H., Mohammed-Geba, K., & Garcia-Vazquez, E. (2017). DNA in a bottle - Rapid metabarcoding survey for early alerts of invasive species in ports. *PLoS ONE*, *12*(9), e0183347. doi: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. doi:10.1007/s10311-016-0581-3

Brandt, A., Gutt, J., Hildebrandt, M., Pawlowski, J., Schwendner, 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. doi:10.3390/jmse4020036

Bucklin, A., Yeh, H. D., Questel, J. M., Richardson, D. E., Reese, B., Copley, N. J., & Wiebe, P. H. (2019). Time-series metabarcoding analysis of zooplankton diversity of the NW Atlantic continental shelf. *ICES Journal of Marine Science*, *76*(4), 1162–1176. doi:10.1093/icesjms/fsz021

Butchart, S. H. M., Walpole, M., Collen, B., Van Strien, A., Scharlemann, J. P. W., Almond, R. E. A., … Watson, R. (2010). Global biodiversity: Indicators of recent declines. *Science*. doi:10.1126/science.1187512

Cadotte, M. W., Dinnage, R., & Tilman, D. (2012). Phylogenetic diversity promotes ecosystem stability. *Ecology*, *93*(8 SPEC. ISSUE), 223–233.

Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME Journal*, *11*(12), 2639–2643. doi:10.1038/ismej.2017.119

Cardinale, B. J., Nelson, K., & Palmer, M. A. (2000). Linking species diversity to the functioning of ecosystems: On the importance of environmental context. *Oikos*, *91*(1), 175–183. doi:10.1034/j.1600-0706.2000.910117.x

Cardinale, Bradley J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., … Naeem, S. (2012). Biodiversity loss and its impact on humanity. *Nature*, *486*(7401), 59–67. doi:10.1038/nature11148

Carignan, V., & Villard, M. A. (2002). Selecting indicator species to monitor ecological integrity: A review. *Environmental Monitoring and Assessment*, *78*(1), 45–61. doi:10.1023/A:1016136723584

Carpenter, S. R., Cole, J. J., Pace, M. L., Batt, R., Brock, W. A., Cline, T., … Weidel, B. (2011). Early warnings of regime shifts: A whole-ecosystem experiment. *Science*, *332*(6033), 1079–1082. doi:10.1126/science.1203672

Carr, C. E., Mojarro, A., Hachey, J., Saboda, K., Tani, J., Bhattaru, S. A., … Ruvkun, G. (2017). Towards in situ sequencing for life detection. *IEEE Aerospace Conference Proceedings*, 1–18. doi:10.1109/AERO.2017.7943896

Cavender-Bares, J., Kozak, K. H., Fine, P. V. A., & Kembel, S. W. (2009). The merging of community ecology and phylogenetic biology. *Ecology Letters*, *12*(7), 693–715. doi:10.1111/j.1461-0248.2009.01314.x

Cavicchioli, R., Ripple, W. J., Timmis, K. N., Azam, F., Bakken, L. R., Baylis, M., … Webster, N. S. (2019). Scientists' warning to humanity: microorganisms and climate change. *Nature Reviews Microbiology*, *17*, 569–586. doi:10.1038/s41579-019-0222-5

CEN. (2018a). *Water quality - Technical report for the routine sampling of benthic diatoms from rivers and lakes adapted for metabarcoding analyses. Technical Report TR17245*.

CEN. (2018b). *Water quality - Technical report for the management of diatom barcodes. Technical Report TR17244.*

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. doi:10.1071/MF15190

Chariton, Anthony A., Court, L. N., Hartley, D. M., Colloff, M. J., & Hardy, C. M. (2010). Ecological assessment of estuarine sediments by pyrosequencing eukaryotic ribosomal DNA. *Frontiers in Ecology and the Environment*, *8*(5), 233–238. doi:10.1890/090115

Chariton, Anthony A., Ho, K. T., Proestou, D., Bik, H., Simpson, S. L., Portis, L. M., … 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. doi:10.1002/etc.2450

Chariton, Anthony 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. doi:10.1016/j.envpol.2015.03.047

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. doi:10.1111/oik.05968

Cheaib, B., Boulch, M. Le, 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. doi: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*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 January 2020 doi:10.20944/preprints202001.0278.v1

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

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. doi:10.1111/mec.13779

Collins, R. A., Wangensteen, 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. doi: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*. doi:10.3389/FEVO.2019.00395

Cordier, T. (2019). Bacterial communities taxonomic and functional turnovers both accurately predict marine benthic ecological quality status. *Environmental DNA*, in press.

Cordier, T., Esling, 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. doi: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. doi:10.1111/1755-0998.12926

Cordier, T., Lanzén, A., Apothéloz-Perret-Gentil, L., Stoeck, T., & Pawlowski, J. (2019a). Embracing environmental genomics and machine learning for routine biomonitoring. *Trends in Microbiology*, *27*(5), 387–397. doi:10.1016/j.tim.2018.10.012

Cordier, T., Frontalini, F., Cermakova, K., Apothéloz-Perret-Gentil, L., Treglia, M., Scantamburlo, E., … Pawlowski, J. (2019b). Multi-marker eDNA metabarcoding survey to assess the environmental impact of three offshore gas platforms in the North Adriatic Sea (Italy). *Marine Environmental Research*, *146*(April 2019), 24–34. doi:10.1016/j.marenvres.2018.12.009

Corse, E., Tougard, C., Archambaud-Suard, G., Agnèse, J. F., Messu Mandeng, F. D., Bilong Bilong, C. F., … Dubut, V. (2019). One-locus-several-primers: A strategy to

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

improve the taxonomic and haplotypic coverage in diet metabarcoding studies. *Ecology and Evolution*. doi:10.1002/ece3.5063

Cougoul, A. P., Bailly, X., & Wit, E. C. (2019). MAGMA: inference of sparse microbial association networks. *BioRxiv*, 538579. doi:10.1101/538579

Creer, S., Fonseca, V. G., Porazinska, D. L., Giblin-Davis, R. M., Sung, W., Power, D. M., … Thomas, W. K. (2010). Ultrasequencing of the meiofaunal biosphere: Practice, pitfalls and promises. *Molecular Ecology*, *19*(SUPPL. 1), 4–20. doi:10.1111/j.1365-294X.2009.04473.x

Crisci, C., Ghattas, B., & Perera, G. (2012). A review of supervised machine learning algorithms and their applications to ecological data. *Ecological Modelling*, *240*, 113–122. doi:http://dx.doi.org/10.1016/j.ecolmodel.2012.03.001

Cristescu, M. E. (2019). Can environmental RNA revolutionize biodiversity science? *Trends in Ecology and Evolution*, *34*(8), 694–697. doi:10.1016/j.tree.2019.05.003

Cristescu, M. E., & Hebert, P. D. N. N. (2018). Uses and misuses of environmental DNA in biodiversity science and conservation. *Annual Review of Ecology, Evolution, and Systematics*, *49*(1), 209–230. doi:10.1146/annurev-ecolsys-110617-062306

Daly, A. J., Baetens, J. M., & De Baets, B. (2018). Ecological diversity: Measuring the unmeasurable. *Mathematics*, *6*(7), 119. doi:10.3390/math6070119

Davies, N., Deck, J., Willis, K., Field, D., Meyer, C., Gilbert, J. A., … Assunta-Sansone, S. (2012). A call for an international network of genomic observatories (GOs). *GigaScience*, *1*(5), 1. doi:10.1186/2047-217X-1-5

de Menezes, A., Clipson, N., & Doyle, E. (2012). Comparative metatranscriptomics reveals widespread community responses during phenanthrene degradation in soil. *Environmental Microbiology*, *14*(9), 2577–2588. doi:10.1111/j.1462-2920.2012.02781.x

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., … Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*. doi:10.1111/mec.14350

Deiner, K., Fronhofer, E. A., Mächler, E., Walser, J. C., & Altermatt, F. (2016). Environmental DNA reveals that rivers are conveyer belts of biodiversity information. *Nature Communications*, *7*, 12544. doi:10.1038/ncomms12544

Delgado-Baquerizo, M., Maestre, F. T., Reich, P. B., Jeffries, T. C., Gaitan, J. J., Encinar, D., … Singh, B. K. (2016). Microbial diversity drives multifunctionality in terrestrial ecosystems. *Nature Communications*, *7*, 1–8. doi:10.1038/ncomms10541

Delmas, E., Besson, M., Brice, M. H., Burkle, L. A., Dalla Riva, G. V., Fortin, M. J., … Poisot, T. (2019). Analysing ecological networks of species interactions. *Biological Reviews*, *94*(1), 16–36. doi:10.1111/brv.12433

Deshpande, Reed, Sullivan, Kerkhof, Beigel, & Wade. (2019). Offline next generation metagenomics sequence analysis using MinION detection software (MINDS). *Genes*, *10*(8), 578. doi:10.3390/genes10080578

Dickie, I. A., Boyer, S., Buckley, H. L., Duncan, R. P., Gardner, P. P., Hogg, I. D., … Weaver, L. (2018). Towards robust and repeatable sampling methods in eDNA-based studies. *Molecular Ecology Resources*. doi:10.1111/1755-0998.12907

Dowle, E., Pochon, X., Keeley, N., & Wood, S. (2015). Assessing the effects of salmon farming seabed enrichment using bacterial community diversity and high-throughput sequencing. *FEMS Microbiology Ecology*, *91*(8), fiv089. doi:http://dx.doi.org/10.1093/femsec/fiv089

Dufrêne, M., & Legendre, P. (1997). Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological Monographs*, *67*(3), 345–366. doi:10.2307/2963459

Dulbecco, R. (1986). A turning point in cancer research: sequencing the human genome. *Science*, *231*(4742), 1055–1056. doi:10.1126/science.3945817

Elbrecht, V., Braukmann, T. W. A., Ivanova, N. V., Prosser, S. W. J., Hajibabaei, M., Wright, M., … Steinke, D. (2019). Validation of COI metabarcoding primers for terrestrial arthropods. *PeerJ*, *7*, e7745. doi:10.7717/peerj.7745

Elbrecht, V., & Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass-sequence relationships with an innovative metabarcoding protocol. *PLoS ONE*, *10*(7), 1–16. doi:10.1371/journal.pone.0130324

Elbrecht, V., Peinert, B., & Leese, F. (2017). Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding. *Ecology and Evolution*, *7*(17), 6918–6926. doi:10.1002/ece3.3192

Elbrecht, V., Vamos, E. E., Meissner, K., Aroviita, J., & Leese, F. (2017). Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution*, *8*(10), 1265–1275. doi:10.1111/2041-210X.12789

Escalas, A., Hale, L., Voordeckers, J. W., Yang, Y., Firestone, M. K., Alvarez-Cohen, L., & Zhou, J. (2019). Microbial functional diversity: From concepts to applications. *Ecology and Evolution*. doi:10.1002/ece3.5670

Fahner, N. A., Shokralla, S., Baird, D. J., & Hajibabaei, M. (2016). Large-scale monitoring of plants through environmental DNA metabarcoding of soil: Recovery, resolution, and annotation of four DNA markers. *PLoS ONE*, *11*(6), e0157505. doi:10.1371/journal.pone.0157505

Falk, N., Reid, T., Skoyles, A., Grgicak-Mannion, A., Drouillard, K., & Weisener, C. G. (2019). Microbial metatranscriptomic investigations across contaminant gradients of the Detroit River. *Science of The Total Environment*, *690*, 121–131. doi:10.1016/j.scitotenv.2019.06.451

Faust, K., Lima-Mendez, G., Lerat, J. S., Sathirapongsasuti, J. F., Knight, R., Huttenhower, C., … Raes, J. (2015). Cross-biome comparison of microbial association networks. *Frontiers in Microbiology*, *6*(OCT), 1–13. doi:10.3389/fmicb.2015.01200

Faust, K., & Raes, J. (2012). Microbial interactions: from networks to models. *Nature Reviews Microbiology*, *10*(8), 538–550. doi:10.1038/nrmicro2832

Faust, K., Sathirapongsasuti, J. F., Izard, J., Segata, N., Gevers, D., Raes, J., & Huttenhower, C. (2012). Microbial co-occurence relationships in the human microbiome. *PLoS Computational Biology*, *8*(7), e1002606. doi:10.1371/journal.pcbi.1002606

Ferrera, I., Giner, C. R., Reñé, A., Camp, J., Massana, R., Gasol, J. M., & Garcés, E. (2016). Evaluation of alternative high-throughput sequencing methodologies for the monitoring of marine picoplanktonic biodiversity based on rRNA gene amplicons. *Frontiers in Marine Science*, *3*(AUG), 147. doi:10.3389/fmars.2016.00147

Field, D., Garrity, G., Gray, T., Morrison, N., Selengut, J., Sterk, P., … Wipat, A. (2008). The minimum information about a genome sequence (MIGS) specification. *Nature Biotechnology*, *26*(5), 541–547. doi:10.1038/nbt1360

Flynn, D. F. B., Mirotchnick, N., Jain, M., Palmer, M. I., & Naeem, S. (2011). Functional and phylogenetic diversity as predictors of biodiversity--ecosystem-function relationships. *Ecology*, *92*(8), 1573–1581. doi:10.1890/10-1245.1

Fordyce, S. L., Ávila-Arcos, M. C., Rasmussen, M., Cappellini, E., Romero-Navarro, J. A., Wales, N., … Gilbert, M. T. P. (2013). Deep sequencing of RNA from ancient maize kernels. *PLoS ONE*, *8*(1), e50961. doi:10.1371/journal.pone.0050961

Friedman, J., & Alm, E. J. (2012). Inferring correlation networks from genomic survey data. *PLoS Computational Biology*, *8*(9), e1002687. doi:10.1371/journal.pcbi.1002687

Frontalini, F., Greco, M., Di Bella, L., Lejzerowicz, F., Reo, E., Caruso, A., … Pawlowski, J. (2018). Assessing the effect of mercury pollution on cultured benthic foraminifera community using morphological and eDNA metabarcoding approaches. *Marine Pollution Bulletin*, *129*(2), 512–524. doi:10.1016/j.marpolbul.2017.10.022

Gagic, V., Bartomeus, I., Jonsson, T., Taylor, A., Winqvist, C., Fischer, C., … Bommarco, R. (2015). Functional identity and diversity of animals predict ecosystem functioning better than species-based indices. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1801), 20142620. doi:10.1098/rspb.2014.2620

Galand, P. E., Lucas, S., Fagervold, S. K., Peru, E., Pruski, A. M., Vétion, G., … Guizien, K. (2016). Disturbance increases microbial community diversity and production in marine sediments. *Frontiers in Microbiology*, *7*(DEC), 1–11. doi:10.3389/fmicb.2016.01950

Galand, P. E., Pereira, O., Hochart, C., Auguet, J. C., & Debroas, D. (2018). A strong link between marine microbial community composition and function challenges the idea of functional redundancy. *ISME Journal*, *12*(10), 2470–2478. doi:10.1038/s41396-018-0158-1

Galand, P. E., Salter, I., & Kalenitchenko, D. (2015). Ecosystem productivity is associated with bacterial phylogenetic distance in surface marine waters. *Molecular Ecology*, *24*(23), 5785–5795. doi:10.1111/mec.13347

Gan, W., Gu, Y., Han, J., Li, C. X., Sun, J., & Liu, P. (2017). Chitosan-modified filter paper for nucleic acid extraction and "in situ PCR" on a thermoplastic microchip. *Analytical Chemistry*, *89*(6), 3568–3575. doi:10.1021/acs.analchem.6b04882

Gardham, S., Hose, G. C., Stephenson, S., & Chariton, A. A. (2014). DNA Metabarcoding Meets Experimental Ecotoxicology: Advancing Knowledge on the Ecological Effects of Copper in Freshwater Ecosystems. *Advances in Ecological Research*, *51*, 79–104. doi:10.1016/B978-0-08-099970-8.00007-5

Gerhard, W. A., & Gunsch, C. K. (2019). Metabarcoding and machine learning analysis of environmental DNA in ballast water arriving to hub ports. *Environment International*, *124*(January), 312–319. doi:10.1016/j.envint.2018.12.038

Gerhold, P., Cahill, J. F., Winter, M., Bartish, I. V., & Prinzing, A. (2015). Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). *Functional Ecology*, *29*(5), 600–614. doi:10.1111/1365-2435.12425

Gerlach, J., Samways, M., & Pryke, J. (2013). Terrestrial invertebrates as bioindicators: An overview of available taxonomic groups. *Journal of Insect Conservation*, *17*(4), 831–850. doi:10.1007/s10841-013-9565-9

Gibson, J. F., Shokralla, S., Curry, C., Baird, D. J., Monk, W. A., King, I., & Hajibabaei, M. (2015). Large-scale biomonitoring of remote and threatened ecosystems via high-throughput sequencing. *PLoS ONE*, *10*(10), e0138432. doi:10.1371/journal.pone.0138432

Gilbert, J. A., Meyer, F., Antonopoulos, D., Balaji, P., Brown, C. T., Brown, C. T., … Stevens, R. (2010). Meeting report: the terabase metagenomics workshop and the vision of an earth microbiome project. *Standards in Genomic Sciences*, *3*(3), 243–248. doi:10.4056/sigs.1433550

Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., … Taberlet, P. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution*, *7*(11), 1299–1307. doi:10.1111/2041-210X.12595

Graham, E. B., Knelman, J. E., Schindlbacher, A., Siciliano, S., Breulmann, M., Yannarell, A., … Nemergut, D. R. (2016). Microbes as engines of ecosystem function: When does community structure enhance predictions of ecosystem processes? *Frontiers in Microbiology*, *7*(FEB), 1–10. doi:10.3389/fmicb.2016.00214

Graham, S. E., Chariton, A. A., & Landis, W. G. (2019). Using Bayesian networks to predict risk to estuary water quality and patterns of benthic environmental DNA in Queensland. *Integrated Environmental Assessment and Management*, *15*(1), 93–111. doi:10.1002/ieam.4091

Gray, C., Baird, D. J., Baumgartner, S., Jacob, U., Jenkins, G. B., O'Gorman, E. J., … Woodward, G. (2014). FORUM: Ecological networks: The missing links in biomonitoring science. *Journal of Applied Ecology*, *51*(5), 1444–1449. doi:10.1111/1365-2664.12300

Grey, E. K., Bernatchez, L., Cassey, P., Deiner, K., Deveney, M., Howland, K. L., … Lodge, D. M. (2018). Effects of sampling effort on biodiversity patterns estimated from environmental DNA metabarcoding surveys. *Scientific Reports*, *8*(May), 8843. doi:10.1038/s41598-018-27048-2

Grizzetti, B., Lanzanova, D., Liquete, C., Reynaud, A., & Cardoso, A. C. (2016). Assessing water ecosystem services for water resource management. *Environmental Science and Policy*, *61*, 194–203. doi:10.1016/j.envsci.2016.04.008

Guardiola, M., Wangensteen, O. S., Taberlet, P., Coissac, E., Uriz, M. J., & Turon, X. (2016). Spatio-temporal monitoring of deep-sea communities using metabarcoding of sediment DNA and RNA. *PeerJ*, *4*, e2807. doi:10.7717/peerj.2807

Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., … Gorsky, G. (2016). Plankton networks driving carbon export in the oligotrophic ocean. *Nature*, *532*(7600), 465–470. doi:10.1038/nature16942

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., … Christen, R. (2013). The Protist Ribosomal reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Research*, *41*(D1), D597–D604. doi:10.1093/nar/gks1160

Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G. A. C., & Baird, D. J. (2011). Environmental barcoding: A next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS ONE*, *6*(4), e17497. doi:10.1371/journal.pone.0017497

Hajibabaei, M., Spall, J. L., Shokralla, S., & van Konynenburg, S. (2012). Assessing biodiversity of a freshwater benthic macroinvertebrate community through non-destructive environmental barcoding of DNA from preservative ethanol. *BMC Ecology*, *12*(1), 28. doi:10.1186/1472-6785-12-28

Hauptmann, A. L., Sicheritz-Pontén, T., Cameron, K. A., Bælum, J., Plichta, D. R., Dalgaard, M., & Stibal, M. (2017). Contamination of the Arctic reflected in microbial metagenomes from the Greenland ice sheet. *Environmental Research Letters*, *12*(7), 074019. doi:10.1088/1748-9326/aa7445

He, X., Sutherland, T. F., Pawlowski, J., & Abbott, C. L. (2019). Responses of foraminifera communities to aquaculture-derived organic enrichment as revealed by environmental DNA metabarcoding. *Molecular Ecology*, *28*(5), 1138–1153. doi:10.1111/mec.15007

He, Z., Zhang, P., Wu, L., Rocha, A. M., Tu, Q., Shi, Z., … Zhou, J. (2018). Microbial functional gene diversity predicts groundwater contamination and ecosystem functioning. *MBio*, *9*(1), 1–15. doi:10.1128/mBio.02435-17

Hector, A., & Bagchi, R. (2007). Biodiversity and ecosystem multifunctionality. *Nature*, *448*(7150), 188–190. doi:10.1038/nature05947

Hemme, C. L., Tu, Q., Shi, Z., Qin, Y., Gao, W., Deng, Y., … Zhou, J. (2015). Comparative metagenomics reveals impact of contaminants on groundwater microbiomes. *Frontiers in Microbiology*, *6*(OCT), 1–12. doi:10.3389/fmicb.2015.01205

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 January 2020 doi:10.20944/preprints202001.0278.v1

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

Hering, D., Borja, A., Jones, J. I., Pont, D., Boets, P., Bouchez, A., … Kelly, M. (2018). Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive. *Water Research*, *138*, 192–205. doi:10.1016/j.watres.2018.03.003

Hodkinson, I. D., & Jackson, J. K. (2005). Terrestrial and aquatic invertebrates as bioindicators for environmental monitoring, with particular reference to mountain ecosystems. *Environmental Management*, *35*(5), 649–666. doi:10.1007/s00267-004-0211-x

Holman, L. E., de Bruyn, M., Creer, S., Carvalho, G., Robidart, J., & Rius, M. (2019). Detection of introduced and resident marine species using environmental DNA metabarcoding of sediment and water. *Scientific Reports*, *9*(1), 11559. doi:10.1038/s41598-019-47899-7

Hooper Chapin III, F.S., Ewel, J.J., Hector, A., Inchausti, P., Lavorel, S., Lawton, J.H., Lodge, D.M., Loreau, M., Naeem, S., Schmid, B., Setälä, H., Symstad, A.J., Vandermeer, J., Wardle, D.A., D. U. (2005). Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecological Monographs 2005*, *75*((1)), 3–35.

Hughes, T. P., Kerry, J. T., Baird, A. H., Connolly, S. R., Dietzel, A., Eakin, C. M., … Torda, G. (2018). Global warming transforms coral reef assemblages. *Nature*. doi:10.1038/s41586-018-0041-2

IPBES. (2019). *Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*. *IPBES Global Assessment Summary for Policymakers* (Vol. May 2019). doi:/10.1590/1676-0611201600010001

Ives, A. R., & Carpenter, S. R. (2007). Stability and diversity of ecosystems. *Science*, *317*(5834), 58–62. doi:10.1126/science.1133258

Jarman, S. N., Berry, O., & Bunce, M. (2018). The value of environmental DNA biobanking for long-term biomonitoring. *Nature Ecology and Evolution*, *2*(8), 1192–1193. doi:10.1038/s41559-018-0614-3

Jeunen, G. J., Knapp, M., Spencer, H. G., Taylor, H. R., Lamare, M. D., Stat, M., … Gemmell, N. J. (2019). Species-level biodiversity assessment using marine environmental DNA metabarcoding requires protocol optimization and standardization. *Ecology and Evolution*, *9*(3), 1323–1335. doi:10.1002/ece3.4843

Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., … Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, *16*(10), 1245–1257. doi:10.1111/ele.12162

Ji, Y., Huotari, T., Roslin, T., Schmidt, N. M., Wang, J., Yu, D. W., & Ovaskainen, O. (2019). SPIKEPIPE: A metagenomic pipeline for the accurate quantification of eukaryotic species occurrences and intraspecific abundance change using DNA barcodes or mitogenomes. *Molecular Ecology Resources*, *00*, 1–12. doi:10.1111/1755-0998.13057

Juul, S., Izquierdo, F., Hurst, A., Dai, X., Wright, A., Kulesha, E., … Turner, D. J. (2015). What's in my pot? Real-time species identification on the MinION. *BioRxiv*, 030742. doi:10.1101/030742

Karimi, B., Maron, P. A., Chemidlin-Prevost Boure, N., Bernard, N., Gilbert, D., & Ranjard, L. (2017). Microbial diversity and ecological networks as indicators of environmental quality. *Environmental Chemistry Letters*, *15*(2), 265–281. doi:10.1007/s10311-017-0614-6

Karimi, B., Meyer, C., Gilbert, D., & Bernard, N. (2016). Air pollution below WHO levels decreases by 40 % the links of terrestrial microbial networks. *Environmental Chemistry Letters*, *14*(4), 467–475. doi:10.1007/s10311-016-0589-8

Karr, J. R. (1999). Defining and measuring river health. *Freshwater Biology*, *41*, 221–234.

Karsenti, E., Acinas, S. G., Bork, P., Bowler, C., De Vargas, C., Raes, J., … Wincker, P. (2011). A holistic approach to marine eco-systems biology. *PLoS Biology*, *9*(10), e1001177. doi:10.1371/journal.pbio.1001177

Keck, F., & Kahlert, M. (2019). Community phylogenetic structure reveals the imprint of dispersal-related dynamics and environmental filtering by nutrient availability in freshwater diatoms. *Scientific Reports*, *9*(1), 11590. doi:10.1038/s41598-019-48125-0

Keeley, N., Pochon, X., & Wood, S. A. (2018). Development and validation of a multi-trophic metabarcoding biotic index for benthic organic enrichment biomonitoring. *Ecological Indicators*, *85*(year 3), 1044–1057. doi:10.1016/j.ecolind.2017.11.014

Keeley, N., Wood, S. A., & Pochon, X. (2018). Development and preliminary validation of a multi-trophic metabarcoding biotic index for monitoring benthic organic enrichment. *Ecological Indicators*, *85*, 1044–1057. doi:10.1016/j.ecolind.2017.11.014

Kelly, M., Boonham, N., Juggins, S., Kille, P., Mann, D. G., Pass, D., … Glover, R. (2018). A DNA based diatom metabarcoding approach for Water Framework Directive classification of rivers. *Environment Agency, Horizon House, Deanery Road, Bristol, BS1 5AH*, (March), 157.

Kelly, M. G., Penny, C. J., & Whitton, B. A. (1995). Comparative performance of benthic diatom indices used to assess river water quality. *Hydrobiologia*, *302*(3), 179–188. doi:10.1007/BF00032108

Kelly, R. P., Port, J. A., Yamahara, K. M., Martone, R. G., Lowell, N., Thomsen, P. F., … Crowder, L. B. (2014). Harnessing DNA to improve environmental management. *Science*. doi:10.1126/science.1251156

Kermarrec, L., Franc, A., Rimet, F., Chaumeil, P., Humbert, J. F., & Bouchez, A. (2013). Next-generation sequencing to inventory taxonomic diversity in eukaryotic communities: A test for freshwater diatoms. *Molecular Ecology Resources*, *13*(4), 607–619. doi:10.1111/1755-0998.12105

Kermarrec, Lenaïg, Franc, A., Rimet, F., Chaumeil, P., Frigerio, J.-M., Humbert, J.-F., & Bouchez, A. (2014). A next-generation sequencing approach to river biomonitoring using benthic diatoms. *Freshwater Science*, *33*(1), 349–363. doi:10.1086/675079

Kisand, V., Valente, A., Lahm, A., Tanet, G., & Lettieri, T. (2012). Phylogenetic and functional metagenomic profiling for assessing microbial biodiversity in environmental monitoring. *PLoS ONE*, *7*(8), e43630. doi:10.1371/journal.pone.0043630

Klunder, L., Lavaleye, M. S. S., Filippidi, A., van Bleijswijk, J. D. L., Reichart, G.-J., van der Veer, H. W., … Mienis, F. (2018). Impact of an artificial structure on the benthic community composition in the southern North Sea: assessed by a morphological and molecular approach. *ICES Journal of Marine Science*, fsy114. doi:10.1093/icesjms/fsy114

Knight, R., Vrbanac, A., Taylor, B. C., Aksenov, A., Callewaert, C., Debelius, J., … Dorrestein, P. C. (2018). Best practices for analysing microbiomes. *Nature Reviews Microbiology*, *16*(7), 410–422. doi:10.1038/s41579-018-0029-9

Kozlov, A. M., Zhang, J., Yilmaz, P., Glöckner, F. O., & Stamatakis, A. (2016). Phylogeny-aware identification and correction of taxonomically mislabeled sequences. *Nucleic Acids Research*, *44*(11), 5022–5033. doi:10.1093/nar/gkw396

Krehenwinkel, H., Wolf, M., Lim, J. Y., Rominger, A. J., Simison, W. B., & Gillespie, R. G. (2017). Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. *Scientific Reports*, *7*(1), 1–12. doi:10.1038/s41598-017-17333-x

Kurtz, Z. D., Müller, C. L., Miraldi, E. R., Littman, D. R., Blaser, M. J., & Bonneau, R. A. (2015). Sparse and Compositionally Robust Inference of Microbial Ecological Networks. *PLoS Computational Biology*, *11*(5), 1–25. doi:10.1371/journal.pcbi.1004226

Lacoursière-Roussel, A., Howland, K., Normandeau, E., Grey, E. K., Archambault, P., Deiner, K., … Bernatchez, L. (2018). eDNA metabarcoding as a new surveillance approach for coastal Arctic biodiversity. *International Journal of Business Innovation and Research*. doi:10.1002/ece3.4213

Lajoie, G., & Kembel, S. W. (2019). Making the most of trait-based approaches for microbial ecology. *Trends in Microbiology*, *27*(10), 814–823. doi:10.1016/j.tim.2019.06.003

Lallias, D., Hiddink, J. G., Fonseca, V. G., Gaspar, J. M., Sung, W., Neill, S. P., … Creer, S. (2015a). Environmental metabarcoding reveals heterogeneous drivers of microbial eukaryote diversity in contrasting estuarine ecosystems. *ISME Journal*, *9*(5), 1208–1221. doi:10.1038/ismej.2014.213

Lallias, D., Hiddink, J. G., Fonseca, V. G., Gaspar, J. M., Sung, W., Neill, S. P., … Creer, S. (2015b). Environmental metabarcoding reveals heterogeneous drivers of microbial eukaryote diversity in contrasting estuarine ecosystems. *ISME Journal*, *9*(5), 1208–1221. doi:10.1038/ismej.2014.213

Langille, M. G. I., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., … Huttenhower, C. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology*, *31*(9), 814–821. doi:10.1038/nbt.2676

Lanzén, A., Lekang, K., Jonassen, I., Thompson, E. M., & Troedsson, C. (2016). High-throughput metabarcoding of eukaryotic diversity for environmental monitoring of offshore oil-drilling activities. *Molecular Ecology*, *25*(17), 4392–4406. doi:10.1111/mec.13761

Lanzén, A., Lekang, K., Jonassen, I., Thompson, E. M., & Troedsson, C. (2017). DNA extraction replicates improve diversity and compositional dissimilarity in metabarcoding of eukaryotes in marine sediments. *PLoS ONE*, *12*(6), 1–18. doi:10.1371/journal.pone.0179443

Laroche, O., Wood, S. A., Tremblay, L. A., Ellis, J. I., Lear, G., & Pochon, X. (2018a). A cross-taxa study using environmental DNA/RNA metabarcoding to measure biological impacts of offshore oil and gas drilling and production operations. *Marine Pollution Bulletin*, *127*, 97–107. doi:10.1016/j.marpolbul.2017.11.042

Laroche, O., Pochon, X., Tremblay, L. A., Ellis, J. I., Lear, G., & Wood, S. A. (2018b). Incorporating molecular-based functional and co-occurrence network properties into benthic marine impact assessments. *FEMS Microbiology Ecology*, *94*(11), 167. doi:10.1093/femsec/fiy167

Laroche, O., Wood, S. A., Tremblay, L. A., Ellis, J. I., Lejzerowicz, F., Pawlowski, J., … Pochon, X. (2016). First evaluation of foraminiferal metabarcoding for monitoring environmental impact from an offshore oil drilling site. *Marine Environmental Research*, *120*, 225–235. doi:10.1016/j.marenvres.2016.08.009

Lau, M. K., Borrett, S. R., Baiser, B., Gotelli, N. J., & Ellison, A. M. (2017). Ecological network metrics: Opportunities for synthesis. *Ecosphere*, *8*(8). doi:10.1002/ecs2.1900

Lawes, J. C., Dafforn, K. A., Clark, G. F., Brown, M. V., & Johnston, E. L. (2017). Multiple stressors in sediments impact adjacent hard substrate habitats and across biological domains. *Science of the Total Environment*, *592*, 295–305. doi:10.1016/j.scitotenv.2017.03.083

Layeghifard, M., Hwang, D. M., & Guttman, D. S. (2017). Disentangling Interactions in the Microbiome: A Network Perspective. *Trends in Microbiology*, *25*(3), 217–228. doi:10.1016/j.tim.2016.11.008

Lear, G., Dopheide, A., Ancion, P., & Lewis, G. D. (2011). A comparison of bacterial, ciliate and macroinvertebrate indicators of stream ecological health. *Aquatic Ecology*, *45*(4), 517–527. doi:10.1007/s10452-011-9372-x

Leese, F., Altermatt, F., Bouchez, A., Ekrem, T., Hering, D., Meissner, K., … Zimmermann, J. (2016). DNAqua-Net: Developing new genetic tools for bioassessment and monitoring of aquatic ecosystems in Europe. *Research Ideas and Outcomes*, *2*, e11321. doi:10.3897/rio.2.e11321

Leese, F., Bouchez, A., Abarenkov, K., Altermatt, F., Borja, Á., Bruce, K., … Weigand, A. M. (2018). Why We Need Sustainable Networks Bridging Countries, Disciplines, Cultures and Generations for Aquatic Biomonitoring 2.0: A Perspective Derived From the DNAqua-Net COST Action. *Advances in Ecological Research*, *in press*, 63–99. doi:10.1016/bs.aecr.2018.01.001

Lejzerowicz, F., Esling, P., Pillet, L., Wilding, T. A., Black, K. D., & Pawlowski, J. (2015). High-throughput sequencing and morphology perform equally well for benthic monitoring of marine ecosystems. *Scientific Reports*, *5*(1), 13932. doi:10.1038/srep13932

Leray, M., Knowlton, N., Ho, S.-L., Nguyen, B. N., & Machida, R. J. (2019). GenBank is a reliable resource for 21st century biodiversity research. *Proceedings of the National Academy of Sciences*, 201911714. doi:10.1073/pnas.1911714116

Levin, L. A., Bett, B. J., Gates, A. R., Heimbach, P., Howe, B. M., Janssen, F., … Weller, R. A. (2019). Global observing needs in the deep ocean. *Frontiers in Marine Science*, *6*(May), 241. doi:10.3389/fmars.2019.241

Li, F., Peng, Y., Fang, W., Altermatt, F., Xie, Y., Yang, J., & Zhang, X. (2018). Application of environmental DNA metabarcoding for predicting anthropogenic pollution in rivers. *Environmental Science and Technology*, *52*(20), 11708–11719. doi:10.1021/acs.est.8b03869

Libbrecht, M. W., & Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nat Rev Genet*, *16*(6), 321–332. doi:10.1038/nrg3920

Lima-Mendez, G., Faust, K., Henry, N., Decelle, J., Colin, S., Carcillo, F., … Velayoudon, D. (2015). Determinants of community structure in the global plankton interactome. *Science*, *348*(6237), 1262073_1-1262073_9. doi:10.1126/science.1262073

Liu, C., Yao, M., Stegen, J. C., Rui, J., Li, J., & Li, X. (2017). Long-term nitrogen addition affects the phylogenetic turnover of soil microbial community responding to moisture pulse. *Scientific Reports*, *7*(1), 17492. doi:10.1038/s41598-017-17736-w

Lobo, J., Shokralla, S., Costa, M. H., Hajibabaei, M., & Costa, F. O. (2017). DNA metabarcoding for high-throughput monitoring of estuarine macrobenthic communities. *Scientific Reports*, *7*(1), 15618. doi:10.1038/s41598-017-15823-6

Loreau, M., & de Mazancourt, C. (2013). Biodiversity and ecosystem stability: A synthesis of underlying mechanisms. *Ecology Letters*, *16*(SUPPL.1), 106–115. doi:10.1111/ele.12073

Louca, S., Jacques, S. M. S., Pires, A. P. F., Leal, J. S., Srivastava, D. S., Parfrey, L. W., … Doebeli, M. (2016). High taxonomic variability despite stable functional structure across microbial communities. *Nature Ecology & Evolution*, *1*(December), 0015. doi:10.1038/s41559-016-0015

Louca, S., Parfrey, L. W., & Doebeli, M. (2016). Decoupling function and taxonomy in the global ocean microbiome. *Science*, *353*(6305), 1272–1277. doi:10.1126/science.aaf4507

Louca, S., Polz, M. F., Mazel, F., Albright, M. B. N., Huber, J. A., O'Connor, M. I., … Parfrey, L. W. (2018). Function and functional redundancy in microbial systems. *Nature Ecology and Evolution*, *2*(6), 936–943. doi:10.1038/s41559-018-0519-1

Lupatini, M., Suleiman, A. K. A., Jacques, R. J. S., Antoniolli, Z. I., de Siqueira Ferreira, A., Kuramae, E. E., & Roesch, L. F. W. (2014). Network topology reveals high connectance levels and few key microbial genera within soils. *Frontiers in Environmental Science*, *2*(May), 1–11. doi:10.3389/fenvs.2014.00010

Ma, A., Lu, X., Gray, C., Raybould, A., Tamaddoni-Nezhad, A., Woodward, G., & Bohan, D. A. (2019). Ecological networks reveal resilience of agro-ecosystems to changes in farming management. *Nature Ecology and Evolution*, *3*(2), 260–264. doi:10.1038/s41559-018-0757-2

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. doi: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. doi:10.1038/gim.2012.157

Martínez-Santos, M., Lanzén, A., Unda-Calvo, J., Martín, I., Garbisu, 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. doi: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), 830–8. doi:10.1038/ismej.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. doi:10.1016/j.landurbplan.2006.11.005

Maurer, D. O. N., Nguyen, H. A. I., Robertson, G., & Gerlinger, T. O. M. (1999). The Infaunal Trophic Index (ITI): Its suitability for marine environmental monitoring. *Ecological Applications*, *9*(June 1997), 699–713. doi:10.1890/1051-0761(1999)009[0699:TITIII]2.0.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. doi:10.1111/j.1461-0248.2010.01509.x

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 January 2020 doi:10.20944/preprints202001.0278.v1

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

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

McCann, K. S. (2000). The diversity–stability debate. *Nature*, *405*(May), 228–233.

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. doi: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. doi: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. doi:10.3389/fevo.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. doi: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 approach in the bioassessment of Portuguese rivers using diatoms. *Ecological Indicators*. doi:10.1016/j.ecolind.2019.105470

Mouillot, D., Graham, N. A. J., Villéger, S., Mason, N. W. H., & Bellwood, D. R. (2013). A functional approach reveals community responses to disturbances. *Trends in Ecology and Evolution*, *28*(3), 167–177. doi:10.1016/j.tree.2012.10.004

Mukherjee, A., Chettri, B., Langpoklakpam, J. S., Basak, P., Prasad, A., Mukherjee, A. K., … Chattopadhyay, D. (2017). Bioinformatic approaches including predictive metagenomic profiling reveal characteristics of bacterial response to petroleum hydrocarbon contamination in diverse environments. *Scientific Reports*, *7*(1), 1108. doi:10.1038/s41598-017-01126-3

Nascimento, F. J. A., Lallias, D., Bik, H. M., & Creer, S. (2018). Sample size effects on the assessment of eukaryotic diversity and community structure in aquatic sediments

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

using high-throughput sequencing. *Scientific Reports*, *8*(1), 11737. doi: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. doi:10.1016/j.marpolbul.2017.10.039

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

Nilsson, L. K. J., Sharma, A., Bhatnagar, R. K., Bertilsson, S., & Terenius, O. (2018). Presence of Aedes and Anopheles mosquito larvae is correlated to bacteria found in domestic water-storage containers. *FEMS Microbiology Ecology*, *94*(6), fiy058. doi: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. doi:10.1093/nar/gky1022

Obi, C. C., Adebusoye, S. A., Ugoji, E. O., Ilori, M. O., Amund, 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. doi: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. doi: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. doi: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. doi:https://doi.org/10.1101/642199

Pawlowski, J., Esling, P., Lejzerowicz, F., Cedhagen, T., & Wilding, T. A. (2014). Environmental monitoring through protist next-generation sequencing metabarcoding:

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

Assessing the impact of fish farming on benthic foraminifera communities. *Molecular Ecology Resources*, *14*(6), 1129–1140. doi:10.1111/1755-0998.12261

Pawlowski, J., Esling, P., Lejzerowicz, F., Cordier, T., Visco, J. A., Martins, C. I. M., … Cedhagen, T. (2016). Benthic monitoring of salmon farms in Norway using foraminiferal metabarcoding. *Aquaculture Environment Interactions*, *8*, 371–386.

Pawlowski, J., Kelly-Quinn, M., Altermatt, F., Apothéloz-Perret-Gentil, L., Beja, P., Boggero, A., … Kahlert, 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. doi:10.1016/j.scitotenv.2018.05.002

Pawlowski, J., Lejzerowicz, F., & Esling, P. (2014). Next-generation environmental diversity surveys of foraminifera: Preparing the future. *Biological Bulletin*.

Payne, R. J. (2013). Seven reasons why protists make useful bioindicators. *Acta Protozoologica*, *52*(3), 105–113. doi:10.4467/16890027AP.13.0011.1108

Pearman, J. K., Aylagas, E., Voolstra, C. R., Anlauf, H., Villalobos, R., & Carvalho, S. (2019). Disentangling the complex microbial community of coral reefs using standardized Autonomous Reef Monitoring Structures (ARMS). *Molecular Ecology*. doi:10.1111/mec.15167

Pellissier, L., Albouy, C., Bascompte, J., Farwig, N., Graham, C., Loreau, M., … Gravel, D. (2018). Comparing species interaction networks along environmental gradients. *Biological Reviews*, *93*(2), 785–800. doi:10.1111/brv.12366

Pérez-Valera, E., Goberna, M., Faust, K., Raes, J., García, C., & Verdú, M. (2017). Fire modifies the phylogenetic structure of soil bacterial co-occurrence networks. *Environmental Microbiology*, *19*(1), 317–327. doi:10.1111/1462-2920.13609

Piñol, J., Mir, G., Gomez-Polo, P., & Agustí, N. (2015). Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular Ecology Resources*, *15*(4), 819–830. doi:10.1111/1755-0998.12355

Pitsch, G., Bruni, E. P., Forster, D., Qu, Z., Sonntag, B., Stoeck, T., & Posch, T. (2019). Seasonality of planktonic freshwater ciliates: Are analyses based on V9 regions of the 18S rRNA gene correlated with morphospecies counts? *Frontiers in Microbiology*, *10*(FEB), 1–15. doi:10.3389/fmicb.2019.00248

Pochon, X., Wood, S. A., Keeley, N. B., Lejzerowicz, F., Esling, P., Drew, J., & Pawlowski, J. (2015). Accurate assessment of the impact of salmon farming on benthic sediment

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

enrichment using foraminiferal metabarcoding. *Marine Pollution Bulletin*, *100*(1), 370–382. doi:10.1016/j.marpolbul.2015.08.022

Pochon, Xavier, Zaiko, A., Fletcher, L. M., Laroche, O., & Wood, S. A. (2017). Wanted dead or alive? Using metabarcoding of environmental DNA and RNA to distinguish living assemblages for biosecurity applications. *PLoS ONE*, *12*(11), 1–19. doi:10.1371/journal.pone.0187636

Poikane, S., Zampoukas, N., Borja, A., Davies, S. P., van de Bund, W., & Birk, S. (2014). Intercalibration of aquatic ecological assessment methods in the European Union: Lessons learned and way forward. *Environmental Science and Policy*, *44*, 237–246. doi:10.1016/j.envsci.2014.08.006

Pollock, J., Glendinning, L., Wisedchanwet, T., & Watson, M. (2018). The madness of microbiome : attempting to find consensus. *Applied and Environmental Microbiology*, *84*(7), 1–12.

Pont, D., Hugueny, B., Beier, U., Goffaux, D., Melcher, A., Noble, R., … Schmutz, S. (2006). Assessing river biotic condition at a continental scale: A European approach using functional metrics and fish assemblages. *Journal of Applied Ecology*, *43*(1), 70–80. doi:10.1111/j.1365-2664.2005.01126.x

Prygiel, J., & Coste, M. (2000). Guide méthodologique pour la mise en oeuvre de l'Indice Biologique Diatomées NF T 90-354. *Agences de l'Eau, Ministere de l'Aménagement Du Territoire et de l'Environnement, Direction de l'Eau & CEMAGREF, Paris, France.*

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., … Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *41*, D590–D596. doi:10.1093/nar/gks1219

Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., & Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nature Biotechnology*, *35*(12), 1211. doi:10.1038/nbt1217-1211b

Quinn, R. A., Navas-Molina, J. A., Hyde, E. R., Song, S. J., Vázquez-Baeza, Y., Humphrey, G., … Knight, R. (2016). From sample to multi-omics conclusions in under 48 hours. *MSystems*, *1*(2). doi:10.1128/msystems.00038-16

Raes, J., Letunic, I., Yamada, T., Jensen, L. J., & Bork, P. (2011). Toward molecular trait-based ecology through integration of biogeochemical, geographical and metagenomic data. *Molecular Systems Biology*, *7*(473), 1–9. doi:10.1038/msb.2011.6

Ramond, P., Sourisseau, M., Simon, N., Romac, S., Schmitt, S., Rigaut-Jalabert, F., … Siano, R. (2019). Coupling between taxonomic and functional diversity in protistan coastal communities. *Environmental Microbiology*, *21*(2), 730–749. doi:10.1111/1462-2920.14537

Ratajczak, Z., Carpenter, S. R., Ives, A. R., Kucharik, C. J., Ramiadantsoa, T., Stegner, M. A., … Turner, M. G. (2018). Abrupt change in ecological systems: inference and diagnosis. *Trends in Ecology and Evolution*, *33*(7), 513–526. doi:10.1016/j.tree.2018.04.013

Ratnasingham, S., & Hebert, P. D. N. (2007). BOLD: The barcode of life data system: barcoding. *Molecular Ecology Notes*, *7*(3), 355–364. doi:10.1111/j.1471-8286.2007.01678.x

Reynoldson, T. B., & Metcalfe-Smith, J. L. (1992). An overview of the assessment of aquatic ecosystem health using benthic invertebrates. *Journal of Aquatic Ecosystem Health*, *1*(4), 295–308. doi:10.1007/BF00044171

Rimet, F., Chaumeil, P., Keck, F., Kermarrec, L., Vasselon, V., Kahlert, M., … Bouchez, A. (2016). R-Syst::diatom: An open-access and curated barcode database for diatoms and freshwater monitoring. *Database*, *2016:*, baw016. doi:10.1093/database/baw016

Rivera, S. F., Vasselon, V., Jacquet, S., Bouchez, A., Ariztegui, D., & Rimet, F. (2018). Metabarcoding of lake benthic diatoms: from structure assemblages to ecological assessment. *Hydrobiologia*, *807*(1), 37–51. doi:10.1007/s10750-017-3381-2

Rivera, Sinziana F., Vasselon, V., Bouchez, A., & Rimet, F. (2020). Diatom metabarcoding applied to large scale monitoring networks: Optimization of bioinformatics strategies using Mothur software. *Ecological Indicators*. doi:10.1016/j.ecolind.2019.105775

Roy, S., Coldren, C., Karunamurthy, A., Kip, N. S., Klee, E. W., Lincoln, S. E., … Carter, A. B. (2018). Standards and guidelines for validating next-generation sequencing bioinformatics pipelines: a joint recommendation of the association for molecular pathology and the college of american pathologists. *Journal of Molecular Diagnostics*, *20*(1), 4–27. doi:10.1016/j.jmoldx.2017.11.003

Ruppert, K. M., Kline, R. J., & Rahman, M. S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Global Ecology and Conservation*, *17*, e00547. doi:10.1016/j.gecco.2019.e00547

Rygg, B., & Norling, K. (2013). *Norwegian Sensitivity Index (NSI) for marine macroinvertebrates, and an update of Indicator Species Index (ISI)*. ISBN:978-82-577-6210-0.

Salis, R. K., Bruder, A., Piggott, J. J., Summerfield, T. C., & Matthaei, C. D. (2017). High-throughput amplicon sequencing and stream benthic bacteria: Identifying the best taxonomic level for multiple-stressor research. *Scientific Reports*, *7*, 44657. doi:10.1038/srep44657

Santini, L., Belmaker, J., Costello, M. J., Pereira, H. M., Rossberg, A. G., Schipper, A. M., … Rondinini, C. (2017). Assessing the suitability of diversity metrics to detect biodiversity change. *Biological Conservation*, *213*, 341–350. doi:10.1016/j.biocon.2016.08.024

Shade, A. (2017). Diversity is the question, not the answer. *ISME Journal*, *11*(1), 1–6. doi:10.1038/ismej.2016.118

Sigsgaard, E. E., Nielsen, I. B., Bach, S. S., Lorenzen, E. D., Robinson, D. P., Knudsen, S. W., … Thomsen, P. F. (2016). Population characteristics of a large whale shark aggregation inferred from seawater environmental DNA. *Nature Ecology & Evolution*, *1*(1), 0004. doi:10.1038/s41559-016-0004

Simonin, M., Voss, K. A., Hassett, B. A., Rocca, J. D., Wang, S. Y., Bier, R. L., … Bernhardt, E. S. (2019). In search of microbial indicator taxa: shifts in stream bacterial communities along an urbanization gradient. *Environmental Microbiology*, *21*(10), 3653–3668. doi:10.1111/1462-2920.14694

Singer, E., Wagner, M., & Woyke, T. (2017). Capturing the genetic makeup of the active microbiome in situ. *ISME Journal*, *11*(9), 1949–1963. doi:10.1038/ismej.2017.59

Smith, M. B., Rocha, A. M., Smillie, C. S., Olesen, S. W., Paradis, C., Wu, L., … Hazen, T. C. (2015). Natural bacterial communities serve as quantitative geochemical biosensors. *MBio*, *6*(3), 1–13. doi:10.1128/mBio.00326-15

Srivastava, D. S., Cadotte, M. W., Macdonald, A. A. M., Marushia, R. G., & Mirotchnick, N. (2012). Phylogenetic diversity and the functioning of ecosystems. *Ecology Letters*, *15*(7), 637–648. doi:10.1111/j.1461-0248.2012.01795.x

Stark, J. D. (1998). SQMCI: A biotic index for freshwater macroinvertebrate coded-abundance data. *New Zealand Journal of Marine and Freshwater Research*, *32*(1), 55–66. doi:10.1080/00288330.1998.9516805

Stark, Z., Dolman, L., Manolio, T. A., Ozenberger, B., Hill, S. L., Caulfied, M. J., … North, K. N. (2019). Integrating genomics into healthcare: a global responsibility. *American Journal of Human Genetics*, *104*(1), 13–20. doi:10.1016/j.ajhg.2018.11.014

Stat, M., Huggett, M. J., Bernasconi, R., Dibattista, J. D., Berry, T. E., Newman, S. J., … Bunce, M. (2017). Ecosystem biomonitoring with eDNA: Metabarcoding across the tree of life in a tropical marine environment. *Scientific Reports*. doi:10.1038/s41598-017-12501-5

Stefanni, S., Stanković, D., Borme, D., de Olazabal, A., Juretić, T., Pallavicini, A., & Tirelli, V. (2018). Multi-marker metabarcoding approach to study mesozooplankton at basin scale. *Scientific Reports*, *8*(1), 12085. doi:10.1038/s41598-018-30157-7

Stoeck, T., Kochems, R., Forster, D., Lejzerowicz, F., & Pawlowski, J. (2018a). Metabarcoding of benthic ciliate communities shows high potential for environmental monitoring in salmon aquaculture. *Ecological Indicators*, *85*, 153–164. doi:10.1016/j.ecolind.2017.10.041

Stoeck, T., Frühe, L., Forster, D., Cordier, T., Martins, C. I. M., & Pawlowski, J. (2018b). Environmental DNA metabarcoding of benthic bacterial communities indicates the benthic footprint of salmon aquaculture. *Marine Pollution Bulletin*, *127*(November 2017), 139–149. doi:10.1016/j.marpolbul.2017.11.065

Stork, N. E. (2018). How many species of insects and other terrestrial arthropods are there on Earth? *Annual Review of Entomology*, *63*(1), 31–45. doi:10.1146/annurev-ento-020117-043348

Sun, D. L., Jiang, X., Wu, Q. L., & Zhou, N. Y. (2013). Intragenomic heterogeneity of 16S rRNA genes causes overestimation of prokaryotic diversity. *Applied and Environmental Microbiology*, *79*(19), 5962–5969. doi:10.1128/AEM.01282-13

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, *21*(8), 2045–2050. doi:10.1111/j.1365-294X.2012.05470.x

Tackmann, J., Rodrigues, J. F. M., & Mering, C. von. (2018). Rapid inference of direct interactions in large-scale ecological networks from heterogeneous microbial sequencing data. *BioRxiv*, 390195. doi:10.1101/390195

Tamaddoni-Nezhad, A., Milani, G. A., Raybould, A., Muggleton, S., & Bohan, D. A. (2013). Construction and validation of food webs using logic-based machine learning and text mining. *Advances in Ecological Research*, *49*, 225–289. doi:10.1016/B978-0-12-420002-9.00004-4

Tapolczai, K., Keck, F., Bouchez, A., Rimet, F., & Vasselon, V. (2019a). Diatom DNA metabarcoding for biomonitoring : strategies to avoid major taxonomical and bioinformatical biases limiting molecular indices capacities. *Frontiers in Ecology and Evolution*. doi:doi: 10.3389/fevo.2019.00409

Tapolczai, K., Vasselon, V., Bouchez, A., Stenger-Kovács, C., Padisák, J., & Rimet, F. (2019b). The impact of OTU sequence similarity threshold on diatom-based bioassessment: A case study of the rivers of Mayotte (France, Indian Ocean). *Ecology and Evolution*, *9*(1), 166–179. doi:10.1002/ece3.4701

Thompson, M. S. A., Bankier, C., Bell, T., Dumbrell, A. J., Gray, C., Ledger, M. E., … Woodward, G. (2016). Gene-to-ecosystem impacts of a catastrophic pesticide spill: testing a multilevel bioassessment approach in a river ecosystem. *Freshwater Biology*, *61*(12), 2037–2050. doi:10.1111/fwb.12676

Thomsen, P. F., Kielgast, J., Iversen, L. L., Møller, P. R., Rasmussen, M., & Willerslev, E. (2012). Detection of a diverse marine fish fauna using environmental DNA from seawater samples. *PLoS ONE*, *7*(8), e41732. doi:10.1371/journal.pone.0041732

Tilman, D., Reich, P. B., & Knops, J. M. H. (2006). Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature*, *441*(7093), 629–632. doi:10.1038/nature04742

Tkacz, A., Hortala, M., & Poole, P. S. (2018). Absolute quantitation of microbiota abundance in environmental samples. *Microbiome*, *6*(1), 1–13. doi:10.1186/s40168-018-0491-7

Tucker, C. M., Cadotte, M. W., Carvalho, S. B., Jonathan Davies, T., Ferrier, S., Fritz, S. A., … Mazel, F. (2017). A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological Reviews*, *92*(2), 698–715. doi:10.1111/brv.12252

Tylianakis, J. M., & Morris, R. J. (2017). Ecological networks across environmental gradients. *Annual Review of Ecology, Evolution, and Systematics*, *48*(1), annurev-ecolsys-110316-022821. doi:10.1146/annurev-ecolsys-110316-022821

Tylianakis, J. M., Tscharntke, T., & Lewis, O. T. (2007). Habitat modification alters the structure of tropical host-parasitoid food webs. *Nature*, *445*(7124), 202–205. doi:10.1038/nature05429

United Nations. (n.d.). A/RES/70/1: Sustainable Development Knowledge Plataform.

Urzelai, A., Hernández, A. J., & Pastor, J. (2000). Biotic indices based on soil nematode communities for assessing soil quality in terrestrial ecosystems. *Science of the Total Environment*, *247*(2–3), 253–261. doi:10.1016/S0048-9697(99)00494-5

Vacher, C., Tamaddoni-Nezhad, A., Kamenova, S., Peyrard, N., Moalic, Y., Sabbadin, R., … Bohan, D. A. (2016). Learning ecological networks from next-generation sequencing data. In *Advances in Ecological Research* (Vol. 54, pp. 1–39). doi:10.1016/bs.aecr.2015.10.004

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., … Dejean, T. (2016). Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology*, *25*(4), 929–942. doi:10.1111/mec.13428

Vandenkoornhuyse, P., Dufresne, A., Quaiser, A., Gouesbet, G., Binet, F., Francez, A. J., … Couée, I. (2010). Integration of molecular functions at the ecosystemic level: Breakthroughs and future goals of environmental genomics and post-genomics. *Ecology Letters*, *13*(6), 776–791. doi:10.1111/j.1461-0248.2010.01464.x

Vasselon, V., Bouchez, A., Rimet, F., Jacquet, S., Trobajo, R., Corniquel, M., … Domaizon, I. (2018). Avoiding quantification bias in metabarcoding: Application of a cell biovolume correction factor in diatom molecular biomonitoring. *Methods in Ecology and Evolution*, *9*(4), 1060–1069. doi:10.1111/2041-210X.12960

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

Vasselon, V., Domaizon, I., Rimet, F., Kahlert, M., & Bouchez, A. (2017b). Application of high-throughput sequencing (HTS) metabarcoding to diatom biomonitoring: Do DNA extraction methods matter? *Freshwater Science*, *36*(1), 162–177. doi:10.1086/690649

Venail, P. A., & Vives, M. J. (2013). Phylogenetic distance and species richness interactively affect the productivity of bacterial communities. *Ecology*, *94*(11), 2529–2536. doi: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. doi:10.1371/journal.pone.0057923

Visco, J. A., Apothéloz-Perret-Gentil, L., Cordonier, A., Esling, 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. doi:10.1021/es506158m

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

von Ammon, U., Wood, S. A., Laroche, 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. doi: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. doi:10.1038/s41564-018-0156-0

Waters, C. N., Zalasiewicz, J., Summerhayes, C., Barnosky, A. D., Poirier, C., Gałuszka, A., … Wolfe, A. P. (2016). The Anthropocene is functionally and stratigraphically distinct from the Holocene. *Science*, *351*(6269), aad2622. doi: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. doi:10.1146/annurev.ecolsys.33.010802.150448

Weigand, H., Beermann, A. J., Čiampor, 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. doi:10.1016/j.scitotenv.2019.04.247

Welsh, H. H., & Ollivier, L. M. (1998). Stream Amphibians 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. doi: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. doi: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

 doi:10.20944/preprints202001.0278.v1

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

minimum information about any (x) sequence (MIxS) specifications. *Nature Biotechnology*, *29*(5), 415–420. doi:10.1038/nbt.1823

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

Zappelini, C., Karimi, B., Foulon, J., Lacercat-Didier, 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. doi: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. doi:10.1128/mBio.00122-11.Editor

Zhou, X., Li, Y., Liu, S., Yang, Q., Su, X., 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. doi: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. doi:10.1371/journal.pone.0108793

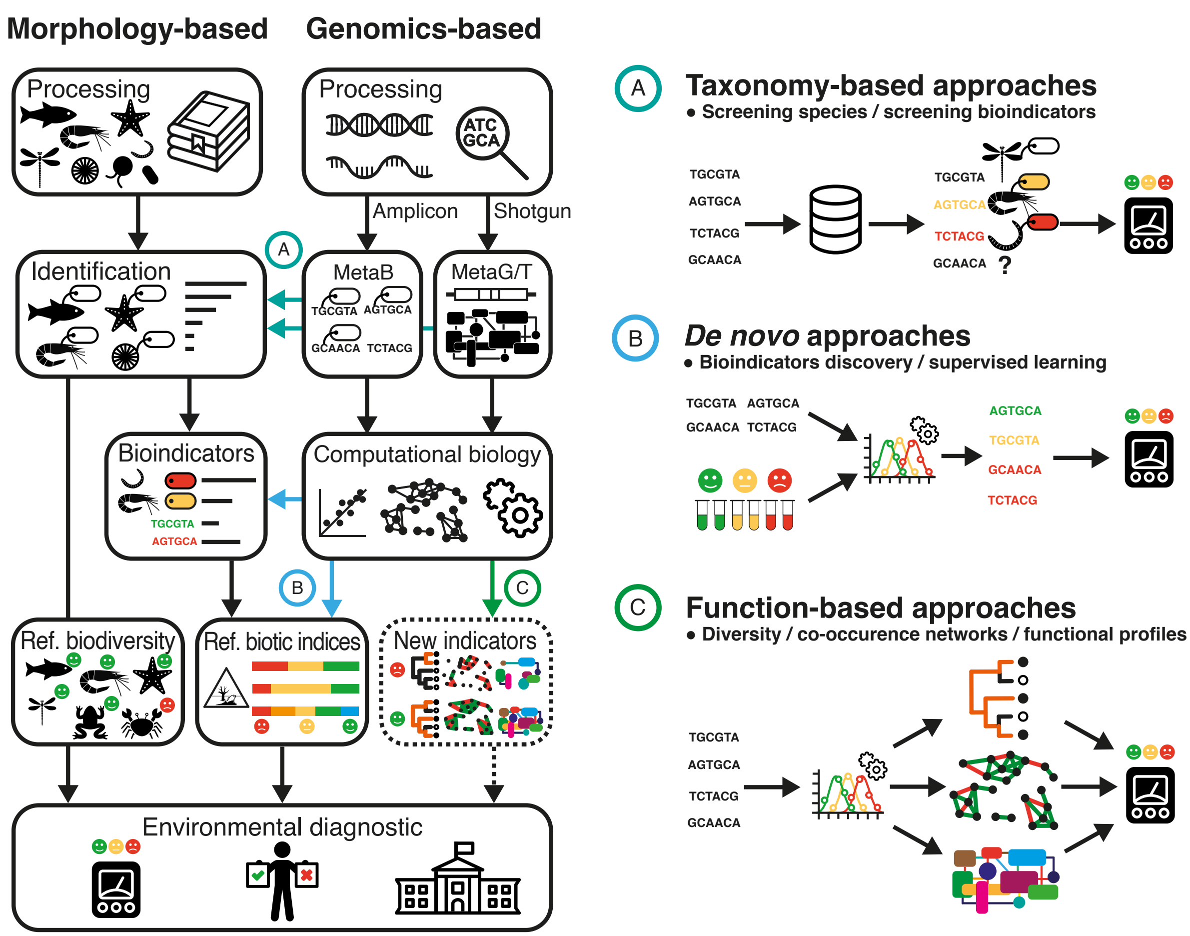
Zinger, L., Bonin, A., Alsos, I. G., Bálint, M., Bik, H., Boyer, F., … Taberlet, P. (2019). DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions. *Molecular Ecology*, *28*(8), 1857–1862. doi:10.1111/mec.15060

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 January 2020 doi:10.20944/preprints202001.0278.v1

Peer-reviewed version available at *Molecular Ecology* **2020**; doi:10.1111/mec.15472

Figures and tables

Figure 1: Conceptual and technical overview of the current three broad analytical strategies environed for ecosystems monitoring with environmental genomics: (A) Taxonomy-based approaches that focus on known bio-indicators or the diversity of taxonomically described taxa, (B) *De novo* approaches that do not require well-established taxonomy, and (C) Function-based approaches that rely on community-wide metrics, where taxa are interchangeable, or on functional profiles instead of compositional turnovers.



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 January 2020 doi:10.20944/preprints202001.0278.v1

Peer-reviewed version available at Molecular Ecology 2020; doi:10.1111/mec.15472

Figure 2: Current envisioned strategy for environmental genomics implementation and their strengths and limitations to fulfill the criteria of existing monitoring programs.

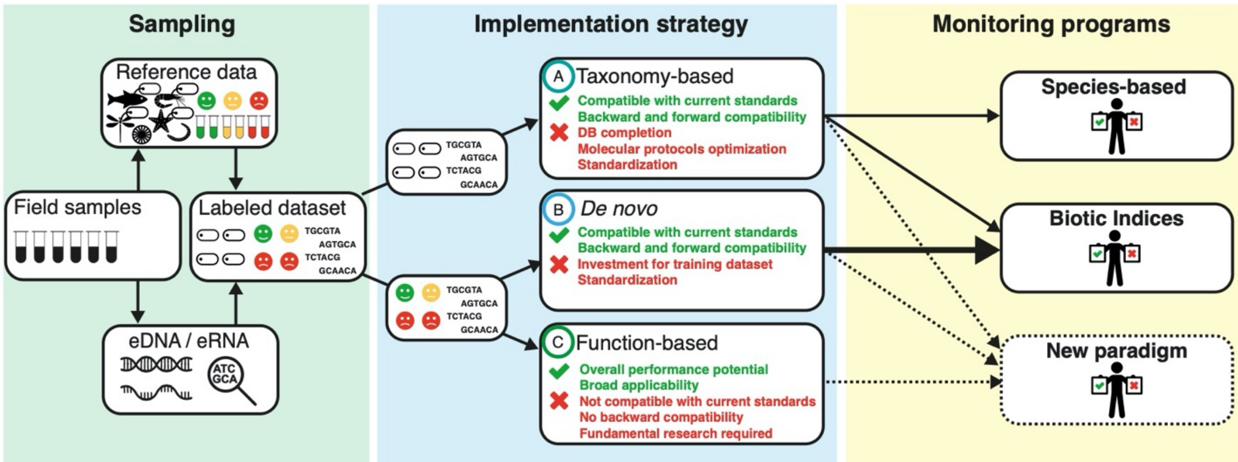


Table 1: List of reviewed studies employing environmental genomics tools for ecosystem monitoring sorted by monitoring strategy, ecosystem, targeted taxonomic group and objective.

Table 2: Comparison of the three strategies in term of compatibility with current standards, backward and forward compatibility, performance, biodiversity coverage, generalization potential and ease of standardization.