For title:

* Components:
  + Comparison metagenomics total rna-seq – that needs to be in title
  + Taxonomic profiling
  + Microbial mock community – that as well
  + Fish tank sample/environmental sample
  + Bioinfo tool benchmarking – relevant for bioinfo people
  + Low sequencing depth – relevant for stakeholders
  + Biomonitoring? – not sure, is not included in any title yet but intro focusses on that. However, what I present doesn’t have to be applied for only biomonitoring

Comparing metagenomics and total RNA sequencing for taxonomic profiling of a microbial mock community and an environmental sample at a moderate sequencing depth using 840 bioinformatic workflows – contains everything, but too long

OR

Comparing metagenomics and total RNA sequencing for taxonomic profiling of microbial communities at a moderate sequencing depth – for stakeholders

OR

Benchmarking of 840 taxonomic profiling workflows for microbial communities using metagenomics and total RNA sequencing – for bioinfo people

OR

Something that’s more specific to biomonitoring, as the intro points in that direction

Intro

1. Freshwater biomonitoring why and how 🡨 Maybe completely exclude the biomonitoring aspect?? Because it’s really long and not really the main focus?
2. Microbial community inclusion was recently advocated
3. Not possible with traditional methods, but DNA based
4. Shotgun sequencing-based approaches, such as metagenomics and metatranscriptomics, present such an option while avoiding primer bias common in metabarcoding
5. But: metagenomics just small percentage of most common microbial barcodes ssu and lsu, and high sequencing depth required 🡪 not sexy for stakeholders
6. In comparison, total RNA-Seq natural enrichment for ssu and lsu, and active portion of community, and low sequencing depth probably enough
7. Comparison of both methods needed

Introduction

PART 1 FRESHWATER ASSESSMENT AND BIOMONITORING, COULD BE (MUCH) SHORTENED IF NECESSARY

Freshwater systems are valuable for economic productivity, ecosystem resilience, and for maintaining ecosystem services, including the supply of clean consumable water (Dudgeon et al., 2006). However, these systems are heavily impacted by agricultural and industrial pollution, habitat fragmentation (wetland drainage, river straightening, and dam building), and the introduction of invasive species (Dextrase and Mandrak, 2006; Jensen et al., 2006). Consequently, freshwater systems belong to the most threatened ecosystems with some of the highest rates of species loss (Malmqvist and Rundle, 2002).

To prevent this species loss, the natural status of freshwater bodies needs to be protected, preserved, and restored. Therefore, the ecological status of individual freshwater bodies must be determined, which is accomplished by freshwater biomonitoring. This determination helps to identify if a freshwater body’s natural status is threatened and if preservation or restoration strategies need to be applied.

Traditionally, freshwater biomonitoring is assessed by sampling bioindicators, most commonly benthic macroinvertebrates, but also diatoms, macrophytes, zooplankton, and fish (Bellinger and Sigee, 2015; Haury et al., 2006; Jeppesen et al., 2011; Karr, 1981; Resh and Unzicker, 1975). However, unicellular eukaryotes and prokaryotes (hereafter referred to as microbes) were proposed as nearly equally good if not better bioindicators, because they respond faster to environmental changes (Ferdous, 2009; Foissner and Berger, 1996; McArthur, 2001; Payne, 2013; Smith et al., 2015; Stoeck et al., 2018). Consequently, recent studies advocate for the inclusion of microbial communities into ecological assessments (Cordier et al., 2019; Pawlowski et al., 2016).

PART 2: DNA-BASED APPROACH

PART 2.1: METABARCODING

For freshwater biomonitoring, sampled communities need to be taxonomically identified. Taxonomic identification of traditional bioindicators is assessed by the morphology of specimens. In contrast to that, taxonomic identification of microbial communities, which is called taxonomic profiling, is often not feasible with morphology. This is due to a lack of diagnostical traits, and DNA-based methods can often be the only option to distinguish microbes (Pawlowski et al., 2012; Will and Rubinoff, 2004). But even for traditional bioindicators, morphological identification can be biased (Stein et al., 2014; Sweeney et al., 2011). Consequently, DNA metabarcoding (Taberlet et al., 2012) has established as a complementing approach in freshwater biomonitoring. It was in fact shown that DNA metabarcoding can even outperform morphology-based ecological assessments (Elbrecht et al., 2017; Stoeck et al., 2018).

Despite this improvement in ecological assessments by DNA metabarcoding, it is known that this approach can be biased as well. Primer bias can be introduced by varying primer binding affinities (Alberdi et al., 2018; Elbrecht and Leese, 2015; Krehenwinkel et al., 2017; Piñol et al., 2014; Piper et al., 2019), and PCR bias can be introduced by GC preference of the polymerase (Nichols et al., 2018; Sze and Schloss, 2019), PCR inhibition (Jane et al., 2015; Taberlet et al., 1996), and chimera formation (Piper et al., 2019; Sze and Schloss, 2019). Furthermore, it is not possible to analyze entire communities with DNA metabarcoding, as primers are designed to target specific groups. Although several studies have shown that targeting multiple groups with multiple primer sets increases taxonomic coverage (Alberdi et al., 2018; De Barba et al., 2014; Stat et al., 2017; Zhang et al., 2018), using multiple primer sets adds bias due to interactions between primers and unequal primer sensitivity. This can lead to varying and unpredictable results (Corse et al., 2019; De Barba et al., 2014; Piñol et al., 2019). Primer interaction may be circumvented by splitting up DNA samples and using different primer sets for every sample (Arulandhu et al., 2017; Swift et al., 2018), but splitting up samples can add stochastic bias, and unequal primer sensitivity is still present.

Both the inability of DNA metabarcoding to target entire communities and the associated primer and PCR bias limit the application of DNA metabarcoding for analyzing microbial communities. These limitations need to be overcome as only unbiased information on entire microbial communities will allow for more accurate ecological assessments.

PART 2.2: METAGENOMICS

An alternative approach for DNA-based taxonomic profiling of microbial communities is metagenomics. This approach allows for taxonomic profiling of entire communities by shotgun sequencing, which involves the random fragmentation and sequencing of the entire DNA in a sample. That way, entire communities can be sequenced PCR-free and, therefore, excluding the bias associated with DNA metabarcoding.

Metagenomics is not commonly used for biomonitoring but has already been applied successfully for years in many other fields (Gilbert and Dupont, 2011; Wooley et al., 2010) and recent studies predict that metagenomics will soon be implemented into biomonitoring (Cordier et al., 2020; Leese et al., 2018). However, the application of metagenomics for taxonomic profiling is highly dependent on sequencing depth, i.e., the sequenced portion of a sample. This dependency is due to the small number of genes in an organism that are useful for taxonomic profiling. Standard barcode genes that are used for that purpose can make up as little as 0.05%-1.4% in metagenomics data (Logares et al., 2014; Yilmaz et al., 2011), and references for other genes are often missing, which is why the major portion of metagenomic data often remains unknown (Singer et al., 2020; Stat et al., 2017). The efficiency of metagenomics for taxonomic profiling is, therefore, dependant on the coverage of these barcode genes, which is directly related to the sequencing depth. The more complex a community is, the higher sequencing depth is required to cover barcode genes to an extend where the community can be accurately profiled. Consequently, given sufficient sequencing depth, metagenomics can outperform metabarcoding in terms of taxonomic profiling (Logares et al., 2014; Shah et al., 2010; Shakya et al., 2013), whereas at low sequencing depths, the coverage of barcodes can sometimes be too low and metagenomics can be outperformed by metabarcoding (Singer et al., 2020; Stat et al., 2017).

Another factor that needs to be considered for metagenomics is the high costs that are involved for high sequencing depths. Given the previously mentioned dependency of metagenomics on sequencing depth, it is recommended to aim for ﻿maximized sequencing output when performing metagenomics. ﻿Illumina HiSeq, NextSeq, and NovaSeq sequencing runs are considered to be suited for that purpose (Quince et al., 2017). However, running these sequencers is expensive, and for many biomonitoring applications, it is important to consider the limited budget available to stakeholders. Consequently, the use of these high sequencing output sequencers might not be affordable yet for routine biomonitoring.

PART 2.3: TOTAL RNA SEQ

Another promising but not broadly applied approach to taxonomic profiling is total RNA sequencing (total RNA-Seq). For this approach, RNA instead of DNA is shotgun sequenced, which is called metatranscriptomics. RNA from a single cell consists to >85% of rRNA and <15% of mRNA (Maaløe, 1979), and therefore, metatranscriptomics usually includes an mRNA enrichment step such as rRNA depletion because the approach is usually focused on mRNA sequencing for gene expression profiling (Bashiardes et al., 2016). However, for total RNA-Seq, no mRNA enrichment step is performed. This enables both the structural and functional analysis of microbial communities (Bang-Andreasen et al., 2020; Li and Guan, 2017; Urich et al., 2008).

Given the high proportion of rRNA in total RNA, total RNA-Seq data is naturally enriched for rRNA sequences and, therefore, for standard barcodes in microbial taxonomic profiling (16S/18S rRNA but also 28S rRNA). For that reason, total RNA-Seq was proposed as a superior alternative to metabarcoding (Geisen et al., 2015) and should also enable better coverage of microbial standard barcodes as compared to metagenomics, which should allow for more accurate taxonomic profiling of complex microbial communities.

Additionally, the natural enrichment in rRNA means that in theory a much lower sequencing depth is required to cover standard barcodes to the same extent as with deep sequencing metagenomics. This translates to reduced sequencing costs for biomonitoring, which is desirable for stakeholders.

The use of RNA instead of DNA for taxonomic profiling has an additional desirable advantage for biomonitoring. Total RNA-Seq can be used to identify the active part of a community (Geisen et al., 2015; Gomez-Silvan et al., 2018), based on the fact that the abundance of rRNA in a cell (as in the number of ribosomes) is an indicator for cell growth and activity (Milo and Phillips, 2015). This signal contrasts with that from metagenomics, which targets the present community, including DNA of dead and/or inactive cells and extracellular DNA, which can make up more than 90% of microbial DNA datasets (Carini et al., 2016; Torti et al., 2015). Information about the active community is desirable for biomonitoring as it reflects the portion of the community that is interacting with the environment and therefore might better reflect environmental conditions.

Comparisons of metagenomics and total RNA-Seq have been carried out for environmental microbial communities (Shi et al., 2011; Yu and Zhang, 2012) and an arthropod mock community (Wilson et al., 2019), but to our knowledge, they have not been compared in taxonomic profiling accuracy by using a microbial mock community.

PART 4: BENCHMARKING

In any sequencing approach, the processing of sequencing data has a high impact on the results. For metagenomics and metatranscriptomics, processing steps include quality filtering, in silico filtering of host DNA/RNA or rRNA, and read- or assembly-based profiling (Knight et al., 2018; Quince et al., 2017), although assembly-based profiling is recommended over read-based profiling (Anwar et al., 2019; Quince et al., 2017). There is an enormous repertoire of processing tools available for each step, and new tools are constantly released. Therefore, benchmarking of available tools is important and frequently carried out, which is thoroughly addressed elsewhere (Bashiardes et al., 2016; Knight et al., 2018; McIntyre et al., 2017; Quince et al., 2017; Shakya et al., 2019; Vollmers et al., 2017). However, to our knowledge, such benchmarking studies have not been performed for total RNA-Seq.

PART 5: WHAT WE DO

In the present study, we target the problem that we are in need for an accurate taxonomic profiling approach for entire diverse microbial communities, yet don’t have the required sequencing depth to do so using metagenomics. Sequencing depth is constantly improving, and sequencing is getting cheaper, but for now, high sequencing depths are still very expensive and often not affordable to stakeholders for routine biomonitoring.

Therefore, we were interested in comparing the performance of metagenomics and total RNA-Seq for taxonomic profiling at a moderate sequencing depth. We chose to compare these two approaches and to exclude DNA metabarcoding due to the limitations associated with the latter. Furthermore, we were interested in the impact of common bioinformatic data processing tools on taxonomic profiles.

To compare the performances of metagenomics and total RNA-Seq and the impact of bioinformatics tools, we applied metagenomics and total RNA-Seq to two sets of samples: 1) a commercially available microbial mock community and 2) a fish tank water sample, which resembles environmental sampling. In addition, we applied XXX (IF WE IMPLEMENT ALL TOOLS ON MY LIST IT WILL BE 840!!! POSSIBLE PIPELINES… MAYBE WE SHOULD REDUCE THAT) different bioinformatic pipelines to both metagenomics and total RNA-Seq data of both sample sets.

Using the known composition of the mock community, we evaluated the accuracy of the sequencing approaches and bioinformatic pipelines in terms of taxa and abundance recovery. This evaluation allowed us to determine the most accurate taxonomic profiling workflow, from sequencing to bioinformatical processing and analysis. We also showed how the taxonomic profile of the fish tank sample differed based on the applied sequencing approaches and bioinformatic pipelines. Finally, we assessed which steps of the bioinformatic pipelines resulted in significant differences in the taxonomic profile for both sample sets and which steps did not have a significant impact.

We predicted that by using total RNA-Seq, we would be able to accurately profile the entire microbial mock community at a moderate sequencing depth and that this approach would outperform metagenomics. This would make total RNA-Seq attractive and affordable for stakeholders. We also predicted that significant differences would be observable for both the microbial mock community and the fish tank sample using different bioinformatic pipelines.

Side note: SSU genes copies can be different within species and organisms (papers roy danzmann QE), that leads to different types of ribosomes in an organism which might cause problems. But same problem exists for metabarcoding and metagenomics and is also never addressed so that should be fine

Methods:

Cutoff for BLAST evalue and for coverage from assemblers

A screenshot of a cell phone

Description automatically generated

Discussion:

Point out that results are valid only for this specific sequencing depth. If sequencing depth increased, metagenomics will become more effective, but we don’t know to what degree/at what depth in relation to community complexity.

RNA hard to handle, short half-life, not feasible for biomonitoring?

Taxonomy-free biomonitoring

Total RNA Seq can be even further used for mRNA analysis, as previously already mentioned – two birds with one stone

Role of mRNA-Seq for taxonomic profiling? (need to find REFS)

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., and Bohmann, K. (2018). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods Ecol. Evol.* 9, 134–147. doi:10.1111/2041-210X.12849.

Anwar, M. Z., Lanzén, A., Bang-Andreasen, T., and Jacobsen, C. S. (2019). To assemble or not to resemble-A validated Comparative Metatranscriptomics Workflow (CoMW). *Gigascience* 8, 1–10. doi:10.1093/gigascience/giz096.

Arulandhu, A. J., Staats, M., Hagelaar, R., Voorhuijzen, M. M., Prins, T. W., Scholtens, I., et al. (2017). Development and validation of a multi-locus DNA metabarcoding method to identify endangered species in complex samples. *Gigascience* 6, 1–18. doi:10.1093/gigascience/gix080.

Bang-Andreasen, T., Anwar, M. Z., Lanzén, A., Kjøller, R., Rønn, R., Ekelund, F., et al. (2020). Total RNA sequencing reveals multilevel microbial community changes and functional responses to wood ash application in agricultural and forest soil. *FEMS Microbiol. Ecol.* 96, 1–13. doi:10.1093/femsec/fiaa016.

Bashiardes, S., Zilberman-Schapira, G., and Elinav, E. (2016). Use of metatranscriptomics in microbiome research. *Bioinform. Biol. Insights* 10, 19–25. doi:10.4137/BBI.S34610.

Bellinger, E. G., and Sigee, D. C. (2015). *Freshwater Algae - Identification, Enumeration and Use as Bioindicators*. 2nd ed. Chichester, West Sussex: John Wiley & Sons Ltd.

Carini, P., Marsden, P. J., Leff, J. W., Morgan, E. E., Strickland, M. S., and Fierer, N. (2016). Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. *Nat. Microbiol.* 2, 1–6. doi:10.1038/nmicrobiol.2016.242.

Cordier, T., Lanzén, A., Apothéloz-Perret-Gentil, L., Stoeck, T., and Pawlowski, J. (2019). Embracing Environmental Genomics and Machine Learning for Routine Biomonitoring. *Trends Microbiol.* 27, 387–397. doi:10.1016/j.tim.2018.10.012.

Cordier, T., Sáez, L. A., Apotheloz-Perret-Gentil, L., Aylagas, E., Bohan, D. A., Bouchez, A., et al. (2020). Ecosystems Monitoring Powered by Environmental Genomics: A Review of Current Strategies with An Implementation Roadmap. *Preprints*. doi:10.20944/PREPRINTS202001.0278.V1.

Corse, E., Tougard, C., Archambaud-Suard, G., Agnèse, J. F., Messu Mandeng, F. D., Bilong Bilong, C. F., et al. (2019). One-locus-several-primers: A strategy to improve the taxonomic and haplotypic coverage in diet metabarcoding studies. *Ecol. Evol.* 9, 4603–4620. doi:10.1002/ece3.5063.

De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E., et al. (2014). DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: Application to omnivorous diet. *Mol. Ecol. Resour.* 14, 306–323. doi:10.1111/1755-0998.12188.

Dextrase, A. J., and Mandrak, N. E. (2006). Impacts of alien invasive species on freshwater fauna at risk in Canada. *Biol. Invasions* 8, 13–24. doi:10.1007/s10530-005-0232-2.

Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z. I., Knowler, D. J., Lévêque, C., et al. (2006). Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biol. Rev. Camb. Philos. Soc.* 81, 163–182. doi:10.1017/S1464793105006950.

Elbrecht, V., and 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, 1–16. doi:10.1371/journal.pone.0130324.

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

Ferdous, Z. (2009). A Review : Potentiality of Zooplankton as Bioindicator. *Am. J. Appl. Sci.* 6, 1815–1819. doi:10.3844/ajassp.2009.1815.1819.

Foissner, W., and Berger, H. (1996). A User-Friendly Guide to the Ciliates. *Freshw. Biol.* 35, 375–482.

Geisen, S., Tveit, A. T., Clark, I. M., Richter, A., Svenning, M. M., Bonkowski, M., et al. (2015). Metatranscriptomic census of active protists in soils. *ISME J.* 9, 2178–2190. doi:10.1038/ismej.2015.30.

Gilbert, J. A., and Dupont, C. L. (2011). Microbial Metagenomics: Beyond the Genome. *Ann. Rev. Mar. Sci.* 3, 347–371. doi:10.1146/annurev-marine-120709-142811.

Gomez-Silvan, C., Leung, M. H. Y., Grue, K. A., Kaur, R., Tong, X., Lee, P. K. H., et al. (2018). A comparison of methods used to unveil the genetic and metabolic pool in the built environment. *Microbiome* 6, 1–16. doi:10.1186/s40168-018-0453-0.

Haury, J., Peltre, M.-C., Trémolières, M., Barbe, J., Thiébaut, G., Bernez, I., et al. (2006). A new method to assess water trophy and organic pollution – the Macrophyte Biological Index for Rivers (IBMR): its application to different types of river and pollution. *Hydrobiologia* 570, 153–158. doi:10.1007/s10750-006-0175-3.

Jane, S. F., Wilcox, T. M., Mckelvey, K. S., Young, M. K., Schwartz, M. K., Lowe, W. H., et al. (2015). Distance, flow and PCR inhibition: EDNA dynamics in two headwater streams. *Mol. Ecol. Resour.* 15, 216–227. doi:10.1111/1755-0998.12285.

Jensen, K., Trepel, M., Merritt, D., and Rosenthal, G. (2006). Restoration ecology of river valleys. *Basic Appl. Ecol.* 7, 383–387. doi:10.1016/j.baae.2006.05.008.

Jeppesen, E., Nõges, P., Davidson, T. A., Haberman, J., Nõges, T., Blank, K., et al. (2011). Zooplankton as indicators in lakes: A scientific-based plea for including zooplankton in the ecological quality assessment of lakes according to the European Water Framework Directive (WFD). *Hydrobiologia* 676, 279–297. doi:10.1007/s10750-011-0831-0.

Karr, J. R. (1981). Assessment of Biotic Integrity Using Fish Communities. *Fisheries* 6, 21–27. doi:10.1577/1548-8446(1981)006.

Knight, R., Navas, J., Quinn, R. A., Sanders, J. G., and Zhu, Q. (2018). Best practices for analysing microbiomes. *Nat. Rev. Microbiol.* doi:10.1038/s41579-018-0029-9.

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

Leese, F., Bouchez, A., Abarenkov, K., Altermatt, F., Borja, Á., Bruce, K., et al. (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. *Adv. Ecol. Res.* 58, 63–99. doi:10.1016/bs.aecr.2018.01.001.

Li, F., and Guan, L. L. (2017). Metatranscriptomic profiling reveals linkages between the active rumen microbiome and feed efficiency in beef cattle. *Appl. Environ. Microbiol.* 83, 1–16. doi:10.1128/AEM.00061-17.

Logares, R., Sunagawa, S., Salazar, G., Cornejo-Castillo, F. M., Ferrera, I., Sarmento, H., et al. (2014). Metagenomic 16S rDNA Illumina tags are a powerful alternative to amplicon sequencing to explore diversity and structure of microbial communities. *Environ. Microbiol.* 16, 2659–2671. doi:10.1111/1462-2920.12250.

Maaløe, O. (1979). “Regulation of the Protein-Synthesizing Machinery---Ribosomes, tRNA, Factors, and So On,” in *Biological Regulation and Development: Gene Expression*, ed. R. F. Goldberger (Boston, MA: Springer US), 487–542. doi:10.1007/978-1-4684-3417-0\_12.

Malmqvist, B., and Rundle, S. (2002). Threats to the running water ecosystems of the world. *Environ. Conserv.* 29, 134–153. doi:10.1017/S0376892902000097.

McArthur, J. V (2001). “Bacteria as Biomonitors,” in *Bioassessment and Management of North American Freshwater Wetlands*, eds. R. B. Rader, D. P. Batzer, and S. A. Wissinger (Chichester, West Sussex: John Wiley & Sons), 249–261. doi:https://doi.org/10.1002/aqc.509.

McIntyre, A. B. R., Ounit, R., Afshinnekoo, E., Prill, R. J., Hénaff, E., Alexander, N., et al. (2017). Comprehensive benchmarking and ensemble approaches for metagenomic classifiers. *Genome Biol.* 18, 1–19. doi:10.1186/s13059-017-1299-7.

Milo, R., and Phillips, R. (2015). *CELL BIOLOGY by the numbers*. Draft. Garland Science,Taylor & Francis Group Available at: http://book.bionumbers.org/what-are-the-rates-of-cytoskeleton-assembly-and-disassembly/.

Nichols, R. V., Vollmers, C., Newsom, L. A., Wang, Y., Heintzman, P. D., Leighton, M., et al. (2018). Minimizing polymerase biases in metabarcoding. *Mol. Ecol. Resour.* 18, 927–939. doi:10.1111/1755-0998.12895.

Pawlowski, J., Audic, S., Adl, S., Bass, D., Belbahri, L., Berney, C., et al. (2012). CBOL Protist Working Group: Barcoding Eukaryotic Richness beyond the Animal, Plant, and Fungal Kingdoms. *PLoS Biol.* 10, e1001419. doi:10.1371/journal.pbio.1001419.

Pawlowski, J., Lejzerowicz, F., Apotheloz-Perret-Gentil, L., Visco, J., and Esling, P. (2016). Protist metabarcoding and environmental biomonitoring: Time for change. *Eur. J. Protistol.* 55, 12–25. doi:10.1016/j.ejop.2016.02.003.

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

Piñol, J., San Andrés, V., Clare, E. L., Mir, G., and Symondson, W. O. C. (2014). A pragmatic approach to the analysis of diets of generalist predators: The use of next-generation sequencing with no blocking probes. *Mol. Ecol. Resour.* 14, 18–26. doi:10.1111/1755-0998.12156.

Piñol, J., Senar, M. A., and Symondson, W. O. C. (2019). The choice of universal primers and the characteristics of the species mixture determine when DNA metabarcoding can be quantitative. *Mol. Ecol.* 28, 407–419. doi:10.1111/mec.14776.

Piper, A. M., Batovska, J., Cogan, N. O. I., Weiss, J., Cunningham, J. P., Rodoni, B. C., et al. (2019). Prospects and challenges of implementing DNA metabarcoding for high-throughput insect surveillance. *Gigascience* 8, 1–22. doi:10.1093/gigascience/giz092.

Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., and Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nat. Biotechnol.* 35, 833–844. doi:10.1038/nbt.3935.

Resh, V. H., and Unzicker, J. D. (1975). Water Quality Monitoring and Aquatic Organisms : The Importance of Species Identification. *Water Pollut. Control Fed.* 47, 9–19.

Shah, N., Tang, H., Doak, T. G., and Ye, Y. (2010). Comparing bacterial communities inferred from 16S rRNA gene sequencing and shotgun metagenomics. *Pacific Symp. Biocomput. 2011*, 165–176.

Shakya, M., Lo, C. C., and Chain, P. S. G. (2019). Advances and challenges in metatranscriptomic analysis. *Front. Genet.* 10, 1–10. doi:10.3389/fgene.2019.00904.

Shakya, M., Quince, C., Campbell, J. H., Yang, Z. K., Schadt, C. W., and Podar, M. (2013). Comparative metagenomic and rRNA microbial diversity characterization using archaeal and bacterial synthetic communities. *Environ. Microbiol.* 15, 1882–1899. doi:10.1111/1462-2920.12086.

Shi, Y., Tyson, G. W., Eppley, J. M., and Delong, E. F. (2011). Integrated metatranscriptomic and metagenomic analyses of stratified microbial assemblages in the open ocean. *ISME J.* 5, 999–1013. doi:10.1038/ismej.2010.189.

Singer, G. A. C., Shekarriz, S., McCarthy, A., Fahner, N., and Hajibabaei, M. (2020). The utility of a metagenomics approach for marine biomonitoring. *bioRxiv*, 2020.03.16.993667. doi:10.1101/2020.03.16.993667.

Smith, M. B., Rocha, A. M., Smillie, C. S., Olesen, S. W., Paradis, C., Wu, L., et al. (2015). Natural Bacterial Communities Serve as Quantitative Geochemical Biosensors. *MBio* 6, e00326-15. doi:10.1128/mBio.00326-15.

Stat, M., Huggett, M. J., Bernasconi, R., Dibattista, J. D., Berry, T. E., Newman, S. J., et al. (2017). Ecosystem biomonitoring with eDNA: Metabarcoding across the tree of life in a tropical marine environment. *Sci. Rep.* 7, 1–11. doi:10.1038/s41598-017-12501-5.

Stein, E. D., White, B. P., Mazor, R. D., Jackson, J. K., Battle, J. M., Miller, P. E., et al. (2014). Does DNA barcoding improve performance of traditional stream bioassessment metrics? *Freshw. Sci.* 33, 302–311. doi:10.1086/674782.

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

Sweeney, B. W., Battle, J. M., Jackson, J. K., and Dapkey, T. (2011). Can DNA barcodes of stream macroinvertebrates improve descriptions of community structure and water quality? *J. North Am. Benthol. Soc.* 30, 195–216. doi:10.1899/10-016.1.

Swift, J. F., Lance, R. F., Guan, X., Britzke, E. R., Lindsay, D. L., and Edwards, C. E. (2018). Multifaceted DNA metabarcoding: Validation of a noninvasive, next-generation approach to studying bat populations. *Evol. Appl.* 11, 1120–1138. doi:10.1111/eva.12644.

Sze, M. A., and Schloss, P. D. (2019). The Impact of DNA Polymerase and Number of Rounds of Amplification in PCR on 16S rRNA Gene Sequence Data. *mSphere* 4, e00163-19. doi:10.1128/msphere.00163-19.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., and Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Mol. Ecol.* 21, 2045–2050. Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1365-294X.2012.05470.x/full%5Cnpapers2://publication/uuid/30F6E470-48F9-4C24-A3C1-6964EE26B34F.

Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., et al. (1996). Reliable genotyping of samll samples with very low DNA quantities using PCR. *Nucleic Acid Res.* 24, 3189–3194.

Torti, A., Lever, M. A., and Jørgensen, B. B. (2015). Origin, dynamics, and implications of extracellular DNA pools in marine sediments. *Mar. Genomics* 24, 185–196. doi:10.1016/j.margen.2015.08.007.

Urich, T., Lanzén, A., Qi, J., Huson, D. H., Schleper, C., and Schuster, S. C. (2008). Simultaneous assessment of soil microbial community structure and function through analysis of the meta-transcriptome. *PLoS One* 3. doi:10.1371/journal.pone.0002527.

Vollmers, J., Wiegand, S., and Kaster, A. K. (2017). *Comparing and evaluating metagenome assembly tools from a microbiologist’s perspective - Not only size matters!* doi:10.1371/journal.pone.0169662.

Will, K. W., and Rubinoff, D. (2004). Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics* 20, 47–55. doi:10.1111/j.1096-0031.2003.00008.x.

Wilson, J.-J., Brandon-Mong, G. J., Gan, H. M., and Sing, K. W. (2019). High-throughput terrestrial biodiversity assessments: mitochondrial metabarcoding, metagenomics or metatranscriptomics? *Mitochondrial DNA Part A* 30, 60–67. doi:10.1080/24701394.2018.1455189.

Wooley, J. C., Godzik, A., and Friedberg, I. (2010). A primer on metagenomics. *PLoS Comput. Biol.* 6. doi:10.1371/journal.pcbi.1000667.

Yilmaz, P., Kottmann, R., Field, D., Knight, R., Cole, J. R., Amaral-Zettler, L., et al. (2011). Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nat. Biotechnol.* 29, 415–420. doi:10.1038/nbt.1823.

Yu, K., and Zhang, T. (2012). Metagenomic and metatranscriptomic analysis of microbial community structure and gene expression of activated sludge. *PLoS One* 7. doi:10.1371/journal.pone.0038183.

Zhang, G. K., Chain, F. J. J., Abbott, C. L., and Cristescu, M. E. (2018). Metabarcoding using multiplexed markers increases species detection in complex zooplankton communities. *Evol. Appl.* 11, 1901–1914. doi:10.1111/eva.12694.