# 2. Advances in DNA-based monitoring to study the effects of marine aggregate extraction on benthic communities

*Derycke Sofie*[*1], *Van Den Bulcke Laure*[1], *Maes Sara*[1], *Lefranc Charles*[1], *Wittoeck Jan*[1], *De Backer Annelies*[1]

[*]*Presenting author: sofie.derycke@ilvo.vlaanderen.be*

[1] *Flanders research institute for agriculture, fisheries and food, ILVO Marine Research, Ankerstraat 1, 8400 Ostend, Belgium*

Marine aggregate extraction activities alter the seafloor through sediment removal and sediment (re)suspension and these seafloor changes in turn affect the benthic communities (de Jong et al., 2015). Understanding whether and how benthic organisms are affected by aggregate extraction provides critical information to safeguard a sustainable exploitation of marine aggregate resources while simultaneously reducing detrimental effects for the marine benthic system. These benthic communities encompass bacteria and archaea, small-sized fauna such as nematodes and copepods (meiobenthos) and animals larger than 1 mm (macrobenthos). In environmental impact assessments (EIAs), benthic metazoan species identification is typically based on morphological characteristics, a time-consuming and labor-intensive process for which specific taxonomic knowledge and experts are needed. DNA-based approaches such as DNA metabarcoding may provide a faster and cheaper alternative to morphological identification.

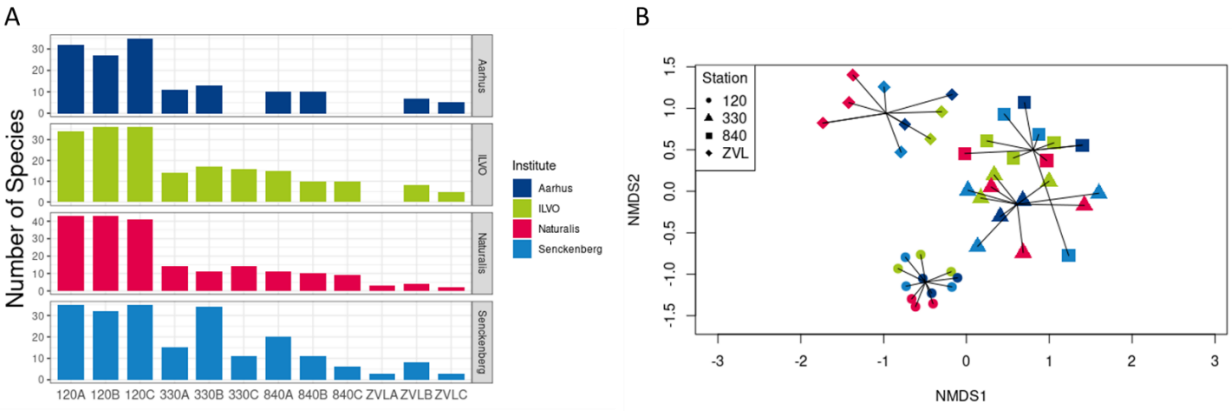
DNA metabarcoding of macrobenthos starts with blending the organisms on a per sample basis to achieve a homogeneous soup from which bulkDNA is extracted. This bulkDNA is used to PCR amplify a portion of the mitochondrial COI gene using general primers and primer specific amplification conditions. The resulting PCR products are sequenced using high throughput sequencing technologies (Baird and Hajibabaei, 2012). Since multiple species in many samples can be sequenced in a single instrument run, processing time of samples is substantially reduced compared to morphological identification (Aylagas et al., 2018). After data analysis, the obtained sequences are linked to species names by comparing them to DNA sequences of morphologically identified specimens in private or public reference databases. Within the North Sea region project GEANS, we tracked time and costs associated with the morphological and DNA-based identification of macrobenthos from the same samples collected in aggregate extraction sites in the Belgian part of the North Sea (BPNS) and showed that DNA-based identification can speed up the identification of macrobenthos samples by 45% while reducing the cost with 27%.

In view of the many laboratory steps in the DNA metabarcoding method, a standardized protocol that allows for reproducible and reliable DNA metabarcoding results is a prerequisite for the adaption of the DNA-based method by policy and stakeholders. Using macrobenthos communities differing in species density and diversity from the BPNS, we first determined the best primer set to PCR amplify as many macrobenthos species as possible using DNA metabarcoding and showed that the DNA-based approach adequately distinguished the different macrobenthos communities (Derycke et al., 2021). Next to the choice of the primer set used to amplify species, the PCR process itself and the DNA extraction step can introduce bias in species detection: during the PCR step, primers may not always adequately bind to the target DNA, while during the DNA extraction step inhibitor substances may be present that affect the PCR efficiency. Therefore, as a second step towards harmonization, we aimed to reduce the stochastic effect of both processes by investigating the number of technical replicates needed in the lab protocol to detect as many species as possible with DNA metabarcoding. Our results showed that three DNA replicates were needed to pick up at least 80% of the species diversity, and at least three PCR replicates in the lab protocol were required to get a good representation of the species (Van den Bulcke et al., 2021). In contrast to general belief, larger body size or higher abundance of the species in a sample did not increase its detection prevalence among DNA replicates. Instead, the diversity in the samples influenced the detection of rare species which were less consistently detected in samples with high diversity compared to locations with less diversity (Van den Bulcke et al., 2021).

Importantly, DNA-based results should be repeatable and robust regardless of the institute that conducts the lab processing of the samples. Therefore, as a third step towards standardization, we conducted a ring test where subsamples of 12 macrobenthos samples originating from four different macrobenthos communities in the BPNS were distributed to four institutes located in Belgium, the Netherlands, Germany and Denmark. Samples were processed using the same standardized lab protocol and the resulting datasets were bioinformatically processed by one institute. Results showed that overall diversity patterns were identical between the four institutes (Fig 1). The number of species showed a similar decreasing trend across institutes from the location with high macrobenthos diversity (station 120) to the station with lowest macrobenthos diversity (ZVL) (Fig 1A). In total, 100 macrobenthos species were detected with DNA metabarcoding, of which 60 species were picked up by all four institutes. At most 14 species were recorded by only one of the four institutes and these species typically had very low abundances. Species composition patterns were also comparable between the four institutes as samples clustered based on the macrobenthos communities independent of the institute that conducted the work (Fig 1B). In addition, small changes to the lab protocol (different DNA extraction kit, different high fidelity polymerases for PCR amplification, different reagents for clean-up) resulted in only minor changes in macrobenthos species detection: similar number of species were detected as with the fixed protocol in all samples and 70 - 75% of the species were shared between the 'fixed' and adjusted protocols. These results show that DNA metabarcoding offers a highly repeatable assessment of species numbers and species composition irrespective of the lab conducting the sample processing.

Figure 1. DNA metabarcoding results of 12 macrobenthos samples that have been processed by four different institutes (indicated by different colors).

*Macrobenthos was collected in station 120 (high diversity, 39 morphological species), stations 330 and 840 (intermediate diversity with 13 and 10 morphological species, respectively) and station ZVL (low diversity with only 3 morphological species). In each station, three biological replicates were collected (A, B, C). A: number of species detected with DNA metabarcoding in each of the 12 stations for each of the four institutes. B: species composition in each of the 12 samples processed by the four institutes illustrated by the nMDS plot based on Bray-Curtis dissimilarity index.*

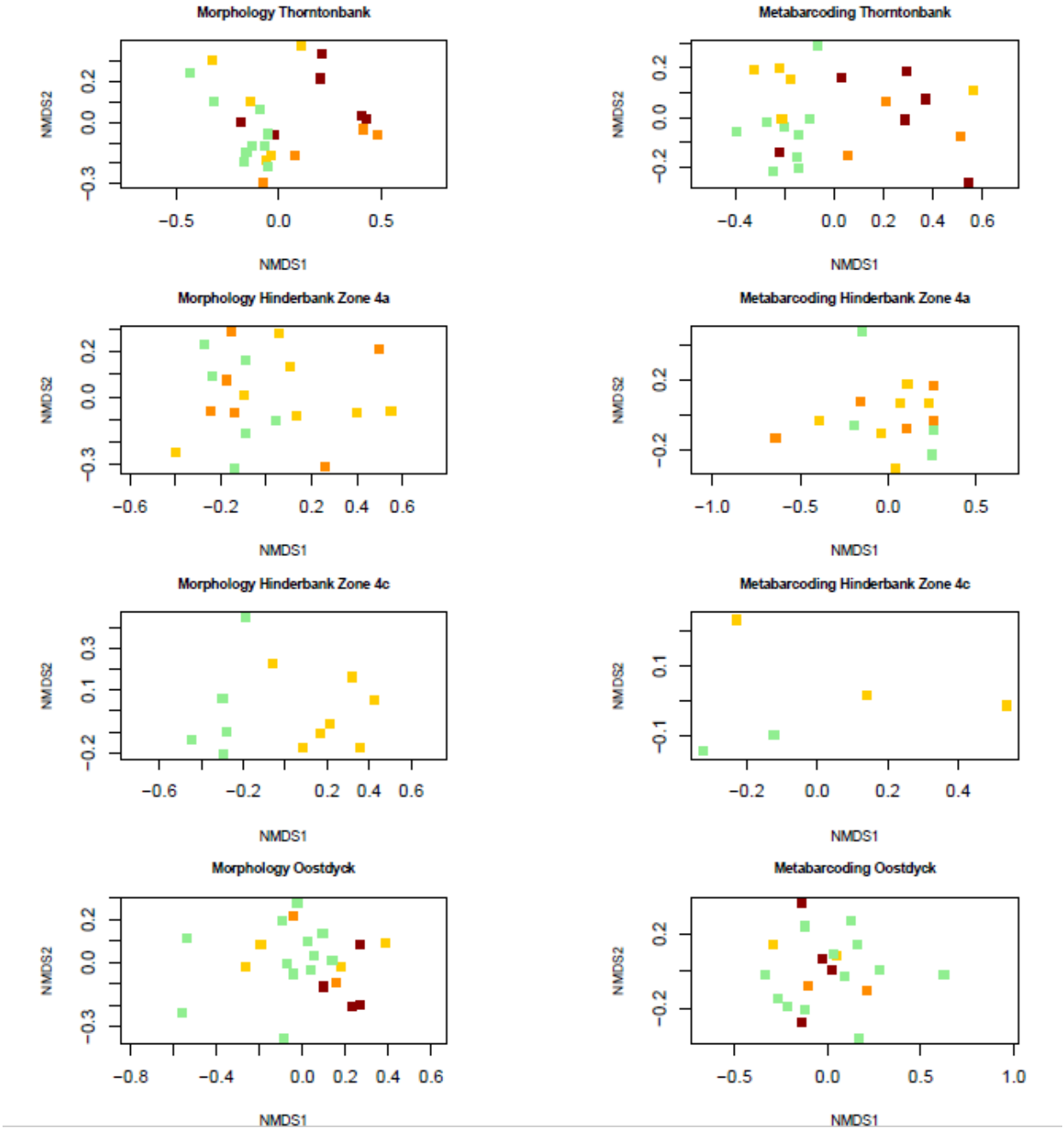



The harmonized and validated DNA metabarcoding protocol was subsequently used to characterize macrobenthos in relation to different regimes of marine aggregate extraction in the BPNS. Macrobenthos samples were collected in 2019 both inside (impact) and outside (reference) extraction areas on three sandbanks characterized by different degrees of extraction intensity: the Thortonbank, which is the epicenter of extraction since 2015 with continuous high extraction intensities of ca 150 000 $m^3$/month and a total extracted volume of 1.8 million $m^3$ in 2019, the Oostdyck with continuous but low extraction intensities of around 30 000 $m^3$/month and a total of 340 000 $m^3$ in 2019 and the Hinderbanken with periodically high amounts of extraction (sometimes up to 500 000 $m^3$/month) for coastal protection, where in 2019 around 600 000 $m^3$ was extracted in zone 4a mainly in the period February-April (Wyns et al., 2021). In zone 4c of the Hinderbanken, no extraction occurred in 2019, but this area has been extracted heavily in previous years. Depending on the amount of sand extracted in 2019, the locations in each of the three sandbanks were divided into three impact groups (high: > 2000 $m^3$, medium: 500 – 2000 $m^3$, low: < 500 $m^3$) and a reference group (0 $m^3$). Multivariate analyses of macrobenthos

communities in the Thortonbank showed a significant impact of sand extraction on species composition for the two methods (PERMANOVA, $p < 0.001$ for DNA metabarcoding and $p = 0.0041$ for morphology). Reference sites were significantly different from the high ($p = 0.024$) , medium ($p = 0.024$) and low ($p = 0.03$) impact sites in the metabarcoding dataset and from the high ($p = 0.012$) and medium ($p = 0.006$) impact sites in the morphological dataset (Fig 2). For the Hinderbanken, the morphological and the metabarcode datasets were consistent and showed no significant differences in macrobenthic communities for zone 4a (Fig 2) suggesting that the recent periodic high extraction in this area has not yet affected the macrobenthos communities. For zone 4c, on the other hand, impact samples clustered separately from reference samples in the morphological and in the metabarcode dataset (Fig 2). No extraction took place in 2019, but previous sand extraction has resulted in finer sediment in the impacted areas compared to the reference areas (Wyns et al., 2021) which may explain the differences in macrobenthos communities between impact and reference samples in zone 4c. Importantly, the number of samples for the metabarcode dataset in zone 4c was reduced to only five because of the low number of DNA sequences obtained for 11 samples. These 11 samples yielded lower DNA and PCR concentrations compared to the other samples despite having a comparable number of specimens, suggesting that DNA quality was reduced and/or PCR amplification was inhibited. This illustrates that the DNA metabarcoding method may not always work for all sample types. For the Oostdyck, species composition was significantly different in the morphological dataset (PERMANOVA, $p = 0.0399$), with high impact sites significantly different from the reference sites ($p = 0.036$, Fig 2). In contrast, no significant differences in species composition were observed for the DNA-based method. This can be explained by the presence of a large number of juveniles in the high impact sites which could not be identified up to species level in the morphological dataset. SIMPER analysis showed that the higher taxon level identifications *Urothoe*, *Bathyporeia*, Echinoidea, *Corophium*, *Nephthys* and *Spio* explained 33% of the differentiation between the high and reference sites. These taxa are regarded as additional species in the morphological dataset thereby artificially inflating species diversity. In the DNA metabarcoding dataset DNA sequences from these juveniles are classified to the correct species, but information on the life stage (juvenile, adult) is lost. These results illustrate that DNA metabarcoding is a valuable method to determine the impact of sand extraction activity on macrobenthos communities and is complementary to morphological identification of macrobenthos samples.

Our next step is to further decrease time and cost associated with sample processing for impact assessment studies by using machine learning algorithms. Machine learning models are trained by using biotic indices based on morphologically identified macrobenthos samples and then use DNA sequence data of macrobenthos to predict the biotic index of new samples using only DNA sequence data. This approach excludes the need for morphological identification, and even for a taxonomic identification step of the DNA sequences thereby circumventing the problem associated with incomplete reference databases for DNA metabarcoding. Furthermore, the machine learning approach can also use DNA sequence data from other organismal groups than macrobenthos (for example bacteria or meiofauna) for which molecular processing time of samples on board and in the lab is much quicker compared to macrobenthos identification. A prerequisite to use other organismal groups to infer the ecological status of samples through machine learning is that they need to show the same response to sand extraction as the macrobenthos. Machine learning algorithms have successfully been used to assess environmental contamination using bacterial communities (Cordier et al., 2019) or foraminifera (Cordier et al., 2017). To explore the potential of machine learning to predict the environmental status of marine aggregate extraction sites, we have conducted 16S rDNA metabarcoding to characterize the bacterial communities in all locations of the three sandbanks described above. We are currently building models to evaluate how well abiotic parameters, COI profiles of macrobenthos and/or 16S profiles of bacterial communities can predict the ecological status of samples determined with morphological macrobenthos data. Our preliminary results indicate that abiotic parameters together with 16S bacterial profiles correlate well with the number of macrobenthos species in sand extraction locations. Additional models are now being built to evaluate whether the prediction of the ecological status of marine aggregate extraction sites can be further improved.

Figure 2: Illustration of macrobenthos species composition for the Thortonbank (top row), Hinderbanken zone 4a (second row), Hinderbanken zone 4c (third row) and Oostdyck (bottom row) using nMDS plots based on the Bray-Curtis dissimilarity index for the morphological dataset (left column) and the DNA metabarcoding dataset (right column). Each square in the plots represents a sample, colors indicate reference sites (green), low impact sites (yellow), medium impact sites (orange) and high impact sites (darkbrown).




In conclusion, our work shows that DNA metabarcoding is a reliable, repeatable and cost-efficient tool for monitoring macrobenthos communities in relation to aggregate extraction in the Belgian part of the North Sea. The DNA-based method decreases time and costs associated with the morphological analyses of macrobenthos samples while adequately capturing changes in macrobenthos diversity. The quick advancements in DNA sequencing technology, bioinformatic processing and machine learning algorithms generate a much higher throughput of samples compared to morphological identification and therefore biodiversity changes related to marine aggregate extraction can be picked up much faster. These methods can be a quick screening and warning tool allowing to use the traditional methods for a more

targeted sampling aiming at understanding the underlying ecological processes and life-history changes. An exciting future for biodiversity monitoring in the marine environment lays ahead of us, with the promise of high resolution biodiversity data generated at unprecedented speed with the sole purpose to achieve a sustainable exploitation of the sea.

**Keywords**: impact assessment, DNA metabarcoding, macrobenthos, biological monitoring, DNA-based monitoring

**Acknowledgements**

This research was supported as part of GEANS (Genetic tools for Ecosystem health Assessment in the North Sea region), an Interreg project supported by the North Sea Program of the European Regional Development Fund of the European Union and co-funding was provided through the sand fund. We thank the crew of RV Belgica and RBINS-OD Nature and BELSPO for providing ship time on RV Belgica.

References:

Aylagas, E., Borja, A., Muxika, I., and Rodriguez-Ezpeleta, N. (2018). Adapting metabarcoding-based benthic biomonitoring into routine marine ecological status assessment networks. *Ecological Indicators* 95, 194-202.

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

Cordier, T., Esling, P., Lejzerowicz, F., Visco, J., Ouadahi, A., Martins, C., Cedhagen, T., and Pawlowski, J. (2017). Predicting the Ecological Quality Status of Marine Environments from eDNA Metabarcoding Data Using Supervised Machine Learning. *Environmental Science & Technology* 51, 9118-9126.

Cordier, T., Lanzen, A., Apotheloz-Perret-Gentil, L., Stoeck, T., and Pawlowski, J. (2019). Embracing Environmental Genomics and Machine Learning for Routine Biomonitoring. *Trends in Microbiology* 27, 387-397.

De Jong, M.F., Baptist, M.J., Lindeboom, H.J., and Hoekstra, P. (2015). Short-term impact of deep sand extraction and ecosystem-based landscaping on macrozoobenthos and sediment characteristics. *Marine Pollution Bulletin* 97, 294-308.

Derycke, S., Maes, S., Van Den Bulcke, L., Vanhollebeke, J., Wittoeck, J., Hillewaert, H., Ampe, B., Haegeman, A., Hostens, K., and De Backer, A. (2021). Detection of Macrobenthos Species With Metabarcoding Is Consistent in Bulk DNA but Dependent on Body Size and Sclerotization in eDNA From the Ethanol Preservative. *Frontiers in Marine Science* 8.

Van Den Bulcke, L., De Backer, A., Ampe, B., Maes, S., Wittoeck, J., Waegeman, W., Hostens, K., and Derycke, S. (2021). Towards harmonization of DNA metabarcoding for monitoring marine macrobenthos: the effect of technical replicates and pooled DNA extractions on species detection. *Metabarcoding and metagenomics* accepted.

Wyns, L., Roche, M., Barette, F., Van Lancker, V., Degrendele, K., Hostens, K., and De Backer, A. (2021). Near-field changes in the seabed and associated macrobenthic communities due to marine aggregate extraction on tidal sandbanks: A spatially explicit bio-physical approach considering geological context and extraction regimes. *Continental Shelf Research* 229.