Received: 20 December 2023 | Revised: 4 April 2024 | Accepted: 9 April 2024

DOI: 10.1002/edn3.552

Environmental DNA
Open Access
Dedicated to the study and use of environmental DNA for basic and applied sciences

WILEY

COMMENTARY

The First International eDNA Workshop in Hong Kong: A beginner's guide for the next-generation eDNA researcher

Mathew Seymour[1] | Isis Guibert[1] | Gert-Jan Jeunen[2] | Micaela Hellström[3] | Matthieu Leray[4] | Vasco Elbrect[5] | Anthony Lau[6] | Maxine Cutracci[1] | Chun Ming How[7] | Jack Chi Ho Ip[8] | Zifang Liu[9] | Jessica Moffitt[2] | Cheng Qian[1] | Xian Qin[7] | Cham Man Tse[1] | Yifu Wang[1] | Zhi Wang[10] | Xiaoyi Zeng[11] | Meihong Zhao[8]

[1]School of Biological Sciences, The University of Hong Kong, Hong Kong SAR, China

[2]Department of Anatomy, School of Biomedical Sciences, University of Otago, Dunedin, New Zealand

[3]MIX Research Sweden AB, Uppsala, Sweden

[4]Smithsonian Tropical Research Institute, Balboa, Ancon, Panama

[5]Independent Researcher and Consultant, Zürich, Switzerland

[6]Lingnan University, Hong Kong SAR, China

[7]City University of Hong Kong, Kowloon, Hong Kong, China

[8]Hong Kong Baptist University, Hong Kong SAR, China

[9]Guangdong University of Technology, Guangzhou, China

[10]State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Xiamen University, Xiamen, China

[11]The Chinese University of Hong Kong, Shenzhen, China

Correspondence
Mathew Seymour, School of Biological Sciences, The University of Hong Kong, Hong Kong SAR, China.
Email: matsey@hku.hk

Funding information
Croucher Foundation Limited (Hong Kong)

Abstract

The field of environmental DNA (eDNA) has rapidly advanced in recent years, providing a non-invasive and time-saving method for assessing biodiversity. The First International Environmental DNA (eDNA) Workshop in Hong Kong was held from 16 to 27 October 2023 and provided early career professionals with hands-on training and collaboration opportunities in eDNA research. With support from The Croucher Foundation Limited (Hong Kong), the workshop covered all stages of an eDNA-based research project, including study design, field sampling, eDNA extraction, library preparation, bioinformatics, statistical data analysis, and ethics in scientific research. Participants gained insights into the principles and practical applications of eDNA technology, emphasizing the importance of careful experimental design and data interpretation. The workshop also highlighted the need for standardized protocols, comprehensive and local DNA reference databases, and careful selection of primer sets to overcome current issues and limitations. Workshop participants expressed enthusiasm for the potential of eDNA metabarcoding as a valuable tool for ecological monitoring, biodiversity assessment, and conservation decision-making. The future of eDNA research looks promising overall, with continued advancements in technology, collaboration among researchers, and the integration of eDNA into large-scale ecological monitoring. Future iterations of the Hong Kong International eDNA workshop will continue to provide hands-on training and collaboration opportunities for early career professionals interested in eDNA research, focusing on addressing current limitations and challenges in the field.

KEYWORDS
biodiversity assessment, bioinformatics, conservation, early career professionals, metabarcoding, study design

1 | BRINGING EDNA KNOWLEDGE EXCHANGE TO HONG KONG

Major developments in scientific understanding come from key needs in human society. We have developed many amazing feats of scientific research in recent years stemming from such needs, including the need to improve our ability to perform complex analytical process by developing artificial intelligence, facilitating direct improvements to our health and longevity through recent advances in medical techniques and pharmaceuticals, and furthering our need to improve our communications and resource acquisition through recent advancements in space technology. We have also made enormous strides in improving our ability to assess and monitoring the living world around us through the recent and ongoing development of environmental DNA (eDNA) based research, which offers a complementary boost to many traditional fields of ecology and environmental sciences. However, as with all scientific advancements, widespread adaptation and continued development of eDNA will depend greatly on our ability to facilitate integration of eDNA into existing fields of research through collaboration and education of new researchers keen to carry on the advancements of eDNA into new areas of study.

Since the conception of eDNA as a means for assessing biological populations and communities, there has been a quick adaptation of eDNA methods to several research fields including ecology, conservation biology, environmental sciences, etc. (Deiner et al., 2017). While more than a decade has passed since the early utilizations of eDNA, rapid development is still ongoing regarding the use and interpretation of eDNA, including adaption of new eDNA sampling methods (Jeunen et al., 2022) to capture lesser observed or dark taxa (Kvist, 2013), optimization of eDNA extraction protocols (Spens et al., 2017), development of optimal biomonitoring pipelines (Hakimzadeh et al., 2023), improvements in taxonomic assignment algorithms (Hleap et al., 2021), and the scrutinization of suitable analytical approaches to convey indirect eDNA data to real world communities (Alberdi & Gilbert, 2019). While advancements in eDNA research will be ongoing for some time, there are now key benchmarks in eDNA methodology that allow for standardized training, particularly for new users that wish to adapt eDNA based methods to their current research.

The First International Environmental DNA (eDNA) Workshop in Hong Kong was hosted by the eDNA & eEcology Lab (Figure 1) and took place from 16 to 27 October 2023 at the University of Hong Kong, Hong Kong SAR, China. Our aim was to provide a set of currently established eDNA practices that would provide new eDNA researchers the opportunity to acquire hands-on experience from established eDNA researchers. Such training was particularly valuable in attracting eDNA research to Southeast Asia where the field is still in the early stages of implementation into mainstream research efforts across academic, industry, and government sectors. In summarizing the efforts of the workshop, this commentary has three key objectives: (1) to summarize the core aspects covered during the 2-week workshop, (2) to highlight key insights regarding the current needs and applications of eDNA research derived from the workshop participants, and (3) to propose future directions for the continued science communication and professional networking of eDNA research.

FIGURE 1 Logo of the hosting lab for the 1st International eDNA workshop in Hong Kong. www.seymourlab.net.

2 | WORKSHOP OVERVIEW

The workshop was financially supported by The Croucher Foundation Limited (croucher.org.hk) and included 30 participants from eight countries and instructors from five nationalities. The target participants of the workshop were early career professionals interested in applying eDNA-based methods to their research efforts. Participants were mostly early career academics, including PhD candidates, postdoctoral fellows, and recently appointed assistant professors. Most of the participants had no prior experience using eDNA, with 25% of the participants having some experience sampling eDNA from water, soil, or as gut microbiome. None of the participants had experience conducting molecular or analytical methods as part of an eDNA study. We also had members from industry and non-profit organizations join as participants. Our instructors were also highly diverse, including researchers from academic institutions, government research units, and private industries.

Over the course of 2 weeks, we introduced participants to all stages of an Edna-based research project. Week 1 included an introduction to eDNA and the fundamentals of study design, onsite field sampling methods, hands-on eDNA extraction laboratory methods, metabarcoding library laboratory preparation, and bioinformatics workflow. Week 2 continued with bioinformatics training, along with taxonomic assignment, data analyses, and data visualization using R and Python programing languages. Throughout the workshop, participants had opportunities to exchange ideas and further develop the direction of the eDNA field and how to promote continued eDNA research networking and knowledge exchange.

3 | INTRODUCTIONS AND GENERAL OVERVIEW

The workshop started with an introduction of the instructors and their background experience in using eDNA to research a wide range of topics, including current method developments of eDNA sampling and extraction (Spens et al., 2017), primer development (Elbrecht & Leese, 2017; Leese et al., 2021) bioinformatics (https://gjeunen.github.io/hku2023eDNAworkshop/bioinformaticanalysis.html), conservation species monitoring (Seymour & Smith, 2023), community assessments (Seymour et al., 2021), and analytical development (Yates et al., 2019).

In highlighting the current state of eDNA research, we focused on the wide breadth of definitions used to describe eDNA, the background of eDNA and its various ecological components, as put forth by Barnes and Turner (2016). There is a considerable amount of debate regarding what eDNA refers to in the academic sphere, such as whether it is restricted to nontargeted multicellular material (Seymour, 2019) or fully inclusive to all molecular material found any given environmental sample (Torti et al., 2015). While the exact definition of what eDNA entails does make for engaging academic discussions, for the sake of simplicity we adopted an inclusive use of the term for the workshop and when referring to the term in this commentary. Subsequently, we focused on four key considerations when using environmental samples as part of an eDNA-based study; (1) first establish the target research questions, objectives and hypotheses, (2) identify the target set of organisms, (3) determine the best study design and whether environmental sampling is best suited, and (4) determine the suitable environmental sample type with regards to the ecology of eDNA.

The final task of the day involved participants forming groups and designing research plans from a set of research questions and scenarios. Participants were keen to tackle the more difficult scenarios, including developing an eDNA-based study to monitoring endangered species. Participants noted the critical need to involve local officials and key stakeholders early in the process, the potential challenges of sampling in difficult or dangerous environments (possibly aided by the use of autonomous vehicles or drones), and the need to establish standard operating procedures (SOP) and eDNA development or practice guidelines as the basis for local or regional standardization. After each presentation, an open discussion among the group was held to discuss logistical and societal considerations and potential alternatives proposed by each group.

4 | FIELD SAMPLING

Surprisingly to many, Hong Kong, with a land area of 1114.35 km^2 possesses a substantial proportion of natural ecosystem coverage. Approximately 40% of the land area (443 km^2) is designated as country parkland with and additional 30% (331 km^2) remaining largely undeveloped, primarily due to the difficult hilly terrain. However, extensive ongoing and planned land reclamation projects threaten to reduce the tangible natural areas of Hong Kong in the coming years (Wang et al., 2021). Additionally, given the high density of people (~7 million) living within Hong Kong, there is an overarching impact on the surrounding wildlife in the form of light pollution, effluent water, and land use changes (Lin et al., 2021).

For the field sampling excursion, we chose a location inside Sai Kung West Country Park that offered convenient access to a costal stream to: (1) showcase standard sediment grab applications along an estuarine site and (2) provide hands on training of eDNA water sampling from coastal water sampling (Figure 2). Here each participant was responsible for processing their own water sample that would be used for the eDNA extraction laboratory practical the following day. The main takeaways we wished to imprint on each of the participants were (1) the need for preplanning the sampling material, (2) maintaining cleanliness throughout the sampling process, and (3) the need to include field collection blanks that are processed as regular samples. When sampling one should also consider that a single point sample will likely miss capturing the total intended area and that multiple samples or subsampling of the targeted area should be utilized, when feasible. As such, taking multiple samples (replicates), and taking multiple subsampling locations, such as sampling and mixing water from along the edge of a lake or a pier is advisable to capture as much of the spatial signal as possible (Bruce et al., 2021). Once samples are collected, and ideally processed in the field, the choice of preservation methods, such as dry storage, ethanol, lysis buffer, or DNA/RNA Shield, is dependent on project logistics and downstream extraction methods.

5 | EDNA EXTRACTION

Prior to laboratory work, participants were introduced to laboratory safety and best practices. Participants were introduced to eDNA extraction protocols and provided hands on laboratory experience, to extract eDNA from their previously collected samples. There are a wide range of protocols for extracting eDNA, including phenol chloroform, ion-exchange, and silica membrane extraction methods. Each extraction method has its own advantages and disadvantages which were discussed and highlighted prior to the laboratory practical. Here we introduced a silica membrane-bases extraction based on the protocol by Spens et al. (2017). This method was chosen as it is widely used and considered safe. However, it does come with a higher cost per sample compared to other extraction methods. In short, the protocol relies on lysis and a proteinase K enzyme washing the membrane that was used to collect the eDNA from the sample. This can be carried out for a couple of hours or overnight. Once the wash has been completed the lysate is collected from the filter and processed following the manufacturer's protocol. Considerations here are to ensure that pipetting skills are practiced beforehand and that filter or plug tips are used throughout the process, to ensure contamination is avoided between samples and from equipment.

6 | METABARCODING

Participants were treated to a series of lectures on primer design and library preparation for Illumina-based high throughput sequencing. The presentation covered the basics of Illumina sequencing methods and terminology, including the use of the MiSeq benchtop sequencer, which can generate up to 30 million paired end reads per run. The presentation also covered the different ways to prepare multiplexed libraries, including ligation, 1-step PCR, and 2-step PCR metabarcoding library pipelines (Bohmann et al., 2022). The advantages and disadvantages of each protocol were discussed, with a focus on amplification bias and hands-on time. Logistics trade-offs

FIGURE 2 Sampling gear and subsequent eDNA samples from the field sampling day. (a) A vertical water sampler used to sample water from a desired depth. (b) Van Veen grab sampler for sediment sampling. (c) Field peristaltic pump. (d) Enclosed filter units, post eDNA sampling.

of each library preparation method were provided for consideration with a consensus that 2-step PCR approaches are largely viewed as the most cost-effective approach. We also discussed recent developments in molecular sequencing platforms, including Oxford Nanopore and Pacific Biosciences which will undoubtably be a factor in future eDNA-based research.

Given the current dominance and standard reliability of Illumina sequencing and the 2-step library preparation approach for eDNA community/biodiversity research, we elected to provide the participants with a hands-on laboratory practical on how to generate Illumina sequencing libraries. As such, participants were provided with a protocol to amplify a commonly used fragment of the COI gene region used in many eDNA research projects (Leray et al., 2013). The main takeaways from the laboratory practical were to gain hands-on practice in standard amplicon library preparation, identify the need to include experimental controls from the field and extraction to identify potential contamination along the process, and to visualize the amplification product via Nanodrop (Thermo Fisher Scientific Inc., Waltham, Massachusetts, United States) and gel-electrophoresis methods. We also provided a demonstration of the Qubit (Thermo Fisher Scientific Inc., Waltham, Massachusetts, United States) quantification, which is commonly used for routine DNA and PCR product quantification and may be more readily available for labs without access to qPCR or fragment analysis machines.

7 | BIOINFORMATICS

eDNA metabarcoding generates a massive amount of data that requires processing via bioinformatic analysis to turn sequencing data into a count or contingency table, suitable for statistical analysis. Bioinformatics involves understanding each step of the bioinformatics workflow and why it is needed. Participants were provided with an extensive practical and series of instruction/discussions covering the full bioinformatics process, which can be found online at https://gjeunen.github.io/hku2023eDNAworkshop/bioinformaticanalysis.html. As most bioinformatics tasks are conducted via terminal commands a basic introduction to Linux/

Unix commands was first provided. This includes operations such as changing directories, creating folders, listing files, and changing permissions that are useful for managing files and folders. Subsequently, participants were informed of the computational constraints and the amount of data lost after each step in a typical high-throughput sequencing bioinformatics pipeline and the associated considerations needed to optimize the workflow. The need for transparency and modular design for building effective bioinformatics workflows was also highlighted with the overarching final output being to generate an operational taxonomic unit (out) amplicon sequence variants (ASV) table, which represents the different taxa present in the samples.

Once sequences have been demultiplexed and quality filtered, bioinformatic processing becomes particularly important in removing noise and artifacts from the sequencing data and to assign taxonomic information to the sequences. There are three main approaches in eDNA metabarcoding workflows: denoizing, clustering, and direct mapping. Each approach has its advantages and limitations. While there are a series of published pipelines available, the key preprocessing steps are largely similar, with notable utilization of the programs Cutadapt (Martin, 2011) for adapter removal and Usearch/Vsearch (Rognes et al., 2016) for sequence searching and clustering. The main takeaways from the bioinformatics were for participants to understand the key steps of the bioinformatics process and to be able to execute each step using a standard pipeline procedure (here using Unix coding). We also provided introduction to the JAMP pipeline (https://github.com/VascoElbrecht/JAMP), an R-based wrapper script approach, as a more a novice-friendly approach to the bioinformatics workflow allowing to analyze the data in a step wise modular fashion.

8 | DATABASE AND TAXONOMY ASSIGNMENT

Taxonomic assignment is a critical step in analyzing metabarcoding data, as it allows researchers to determine the taxonomic identity of the DNA signals present in a sample. This information is essential for understanding community composition, biodiversity, and ecological processes, though taxon-independent methods such as UniFrac distance are also a potential solution (Lozupone et al., 2011). There are several methods and tools available for taxonomic assignment, each with its own advantages and limitations. Common approaches include the use of public databases such as Barcode of Life Data Systems (BOLD) or National Center for Biotechnology Information (NCBI) to obtain reference data for taxonomic assignment. These databases contain a vast amount of genetic information from various organisms, making them valuable open-source resources for researchers. However, the completeness and accuracy of public databases can vary depending on the marker region and primers used in the study. It is important to consider the limitations of any database and validate the results through spot-checking and manual curation. Specialized databases, such as MIDORI2 (Leray et al., 2022) or R-Syst::diatom (Rimet et al., 2016) offer reliable public reference databases and may be applicable for researchers focusing on specific species or geographic area. Another option is to create a local reference database. While this approach requires more effort in setting up and maintaining the database, it offers the advantage of custom taxonomic output and formatting. Databases that are locally curated help ensure accuracy and quality of the utilized reference data. However, local databases may not always be as comprehensive or up to date as public databases and may draw scrutiny if not open access or reproducible. Here, programs such as CRABS offer a hybrid approach to generate reference databased from currently available public data (Jeunen et al., 2023).

When performing taxonomic assignments, it is crucial to spot check the results to verify the accuracy of the assignments. This can be done by manually checking a subset of sequences against known reference sequences or by using tools such as BOLD and NCBI Blast or databases specially designed for certain target groups, such as fish (http://mitofish.aori.u-tokyo.ac.jp/) fungi (https://unite.ut.ee, https://globalfungi.com) or symbiont zooxanthellae (http://www.SymbioGBR.org). Spot checking helps identify potential misidentifications or errors in the taxonomic assignments and allows researchers to correct and refine their analysis.

It is important to note that taxonomic assignment is a separate process from data processing and statistical analysis. Researchers have the flexibility to choose different tools for each step, depending on their specific research questions and requirements. By carefully considering the advantages and limitations of different taxonomic assignment methods and validating the results, researchers can ensure the accuracy and reliability of their metabarcoding data analysis.

9 | DATA ANALYSES AND PRESENTATIONS

Participants were provided with an introduction to R coding and subsequently to a series of exercises on how to generate analyses and figures based on previously published eDNA research. The R code for these exercises is available at https://github.com/MatSeymour/Seymour_etal_eDNA_2024. Given the complexity and time needed, we were unable to provide a full course on statistical analyses but did highlight several key statistical analyses related to community ecology and population ecology data analyses, including linear regression and ordination techniques (Borcard et al., 2018). As most of the interest of the participant group was related to community-based analyses, we also covered the concept of alpha, beta, and gamma biodiversity metrics and how these can be derived from eDNA metabarcoding data. Key takeaways from this portion of the workshop were the expressed need to design the experiment first, including expected statistical test, prior to sampling and data generation. Without properly accounting for the spatial, temporal or treatment replication in the experimental design effective statistical testing to reliably determine differences among samples becomes

essentially impossible leading to loss of scientific integrity, time, and money.

The need for effective scientific illustrations was also highlighted, including the benefits of using illustrations to attract attention in the review and readership phases of academic publishing. Here we provided participants with a crash session on using software such as Inkscape to design illustrations related to study design and diagrams and to the R package ggplot (Wickham, 2016) for illustrating statistical findings from eDNA-derived biodiversity data.

10 | SCIENTIFIC ETHICS

We also highlighted a key need for research within the field of eDNA to be transparent, particularly under modern publishing requirements. A major frustration with published academic research, particularly those funded by public funding bodies, is the lack of data availability and reproducibility of published work. As the field of eDNA is still new, we can set the bar for future and past fields of research by enabling a progressive and open community. Thankfully, eDNA research is already comparably transparent, though far from perfect, given the existing requirements of the genetics field to require molecular data to be made available in public databases. Here we highlighted the need to upload sequence data generated from eDNA studies to local depositories, as well as the need to include final OTU or ASV tables (and associated metadata) with the published paper, preferably as supplementary material.

11 | KEY TAKEAWAYS

During the last day of the workshop, we carried out a participation survey to collect feedback on the effectiveness of the workshop overall. Feedback was provided by 25 out of 30 participants with the survey results summarized as follows. Overall, the workshop was well received by the participants with the majority rating the quality, deliverables, content, and organization as highly relevant or excellent (Figure 3). Additional requests included extending the workshop duration, more in-depth statistical analyses, more information on primer design, and including additional case studies. A comment for improvement was for the bioinformatic and R programing basic to be provided earlier in the workshop or to provide tutorials prior to the start of the workshop. Participants largely commented on the benefits from hands on experience in the field,

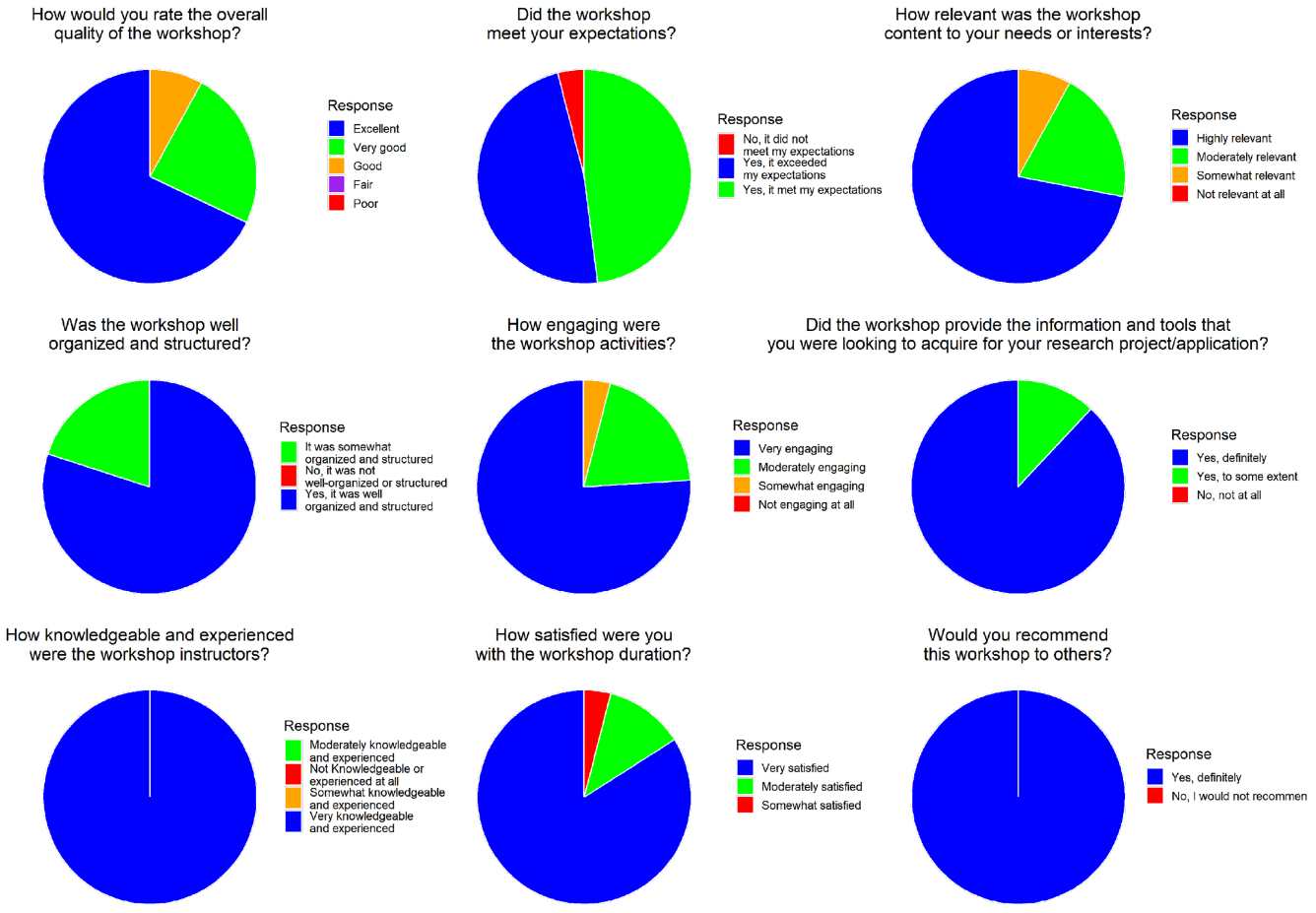

FIGURE 3 Survey results from the 1st International eDNA Workshop. 25 of 30 participants responded to generate each of the subfigures.

lab and coding exercises, with participants noting that a big draw to the workshop was the opportunity to have hands on training over video style instruction.

In writing this commentary, we invited participants (Figure 4) and members of the committee to provide in-depth responses and feedback regarding the current state and direction of the field and on the workshop in general with specific insights provided regarding, (1) the current state and interest of eDNA as a field, (2) overall impressions of the workshop, and (3) key considerations when setting up an eDNA based study. Below are the summaries of the feedback provided.

12 | CURRENT STATE AND INTEREST

The field of environmental DNA (eDNA) has garnered significant interest and attention from various disciplines, across academic, private and industrial sectors. While many seek to utilize eDNA for detecting elusive species, of conservation or management interest, or for biodiversity monitoring and community assessment, the applications of eDNA extends into many other fields, including forensic, agriculture, aquaculture, and natural resource management. Significant advancements have been made in the development and integration of eDNA methods over the past decade, leading to great strides and success. These advancements have played a crucial role in establishing eDNA-based approaches as a standard method for biodiversity detection. As the field of eDNA continues to evolve, there is increasing interest to explore eDNA's potential utility for biomass quantification and integrating it into general biodiversity studies. This integration would unlock its full potential as a comprehensive tool for ecological monitoring. Additionally, there is a growing interest in using eDNA for large-scale ecological monitoring, particularly in marine environments.

Specific eDNA application interest from the participants included assessing macrofauna and meiofauna diversity in marine sediments, studying dinoflagellate composition and distribution, investigating forest soil biodiversity, exploring marine benthic species and their relationships with sediment, and studying the distribution and biodiversity of elasmobranch communities. However, we all acknowledge the challenges associated with the lack of standardized methods, incomplete DNA reference databases, and limited taxonomic resolution at the family or order level for many eDNA studies.

FIGURE 4 Group picture of the workshop participants and instructors.

The development of a comprehensive DNA reference database is crucial to overcoming key limitations in species identification and classification when using eDNA to accurately assess biodiversity.

In terms of method development, participants are keen to make the eDNA workflow more efficient. They are exploring the use of portable machines for DNA extraction, developing efficient PCR protocols, utilizing next-generation sequencing technologies, and integrating advanced algorithms to increase bioinformatic and analytical throughput. The goal is to complete an eDNA-based assessment within hours, compared to days or weeks used for current workflow and months for traditional means, leading to more timely and effective conservation decisions. Additionally, there is a keen interest in looking into examining organisms across multiple trophic levels using third-generation sequencing, which allows for the assessment of entire communities, from bacteria to mammals. Despite these challenges, eDNA has the potential to supplement traditional biodiversity surveys and contribute to effective ecological monitoring and conservation decision-making.

Overall, the participants have a positive outlook on the current state of eDNA and its future direction. They recognize the advancements made in terms of acceptance, guidelines, and research publications. However, they also emphasize the need for standardized protocols, comprehensive DNA reference databases, and taxon-specific primer sets to overcome current limitations. Participants are enthusiastic about the potential of eDNA as a noninvasive, time-saving, and accurate tool for ecological monitoring, biodiversity assessment, and conservation decision-making. With continued research and collaboration, the field of eDNA is expected to further expand its applications and contribute to a deeper understanding of biodiversity and ecosystem dynamics.

13 | WORKSHOP IMPRESSIONS

The eDNA workshop was a comprehensive and well-organized learning experience that provided valuable insights into the principles and practical applications of eDNA technology. From field sampling to bioinformatics analysis, the workshop covered all aspects of eDNA research, emphasizing the importance of careful experimental design and data interpretation. Participants gained knowledge in areas such as preventing potential contamination during collection, employing effective water filtration techniques for sample treatment, determining the necessary replicates for sampling points, setting up a proper eDNA laboratory, and designing appropriate primers. The workshop also highlighted the need for a well-established reference sequence library and the potential for haplotype analysis to provide more quantitative data.

Participants also gained insights into the development of eDNA experiments. The workshop provided an opportunity to think about aspects of planning an eDNA experiment in more depth, particularly regarding considering DNA transportation and pre-collection conditions. This aspect of the workshop significantly increased participants' ability to plan and conduct an eDNA study in a way that uses optimal collection techniques to limit contamination and collection bias, and to consider the extent of ecological inference that can be made from a given sampling design. Bioinformatics was also covered in detail, with participants learning the steps behind bioinformatic pipelines in a way that allows for personalization to specific datasets, something that many had previously struggled with.

The workshop also provided opportunities for collaboration and discussion among participants and instructors. Discussions included experimental setup, primer design and hypothesis formulation. Participants also learned the importance of conducting a pilot study to try out new sampling or molecular methods and to gain a better understanding of the study system and target organism(s). As an emerging field of research, eDNA methods have yet to be as standardized as more traditional methods. Many eDNA studies are often the first to develop or utilize a set of methods, which might need to be modified depending on the location, study system, sample type, etc. As such, carrying out a pilot study to test novel methodology prior to a full study is often beneficial in gaining important information that can save time and resources.

14 | KEY CONSIDERATIONS WHEN SETTING UP AN EDNA-BASED STUDY

When setting up an eDNA study, there are several important considerations to keep in mind. First, it is crucial to have clear research questions and objectives in mind. This will provide a clear direction and purpose for the study. Second, the choice of sample type (e.g. water, soil, and air) is important and should be based on the target phylum's habitat. Different sample types, such as water, sediment, soil, or air samples, may require different collection and preservation methods. It is important to select the most suitable sample type for the research question at hand. It is also important to consider sampling site locations, considering the spatial and temporal scope needed for a given study. Having knowledge on selected sites is important to assess the best sampling methods as well as to avoid potential data drawbacks, such as sample contamination by external sources.

Contamination reduction is another critical aspect to consider in an eDNA study. To minimize contamination, it is important to implement rigorous protocols and use appropriate negative controls at each step of the process, including field sampling, DNA extraction, and PCR stages. Contamination can significantly affect the accuracy and reliability of the results, so careful attention should be given to reducing and monitoring potential sources of contamination.

The choice of primers and the database used for analysis are also vital considerations. Primers should be selected carefully to ensure they are effective in amplifying the target species or taxa with low bias. Depending on the research question, a universal primer or a species-specific primer may be required. Additionally,

the availability and quality of the reference database should be assessed to ensure adequate coverage of target and closely related species. While universal databases provide a quick overview, local or habitat-specific databases can offer more detailed and accurate biodiversity information. It is important to evaluate the suitability of the existing databases or consider creating a suitable database if necessary.

Data analysis is another important consideration in eDNA studies. To ensure accurate analysis and interpretation of the data, eDNA researchers should possess a comprehensive understanding of both biological and statistical concepts or collaborate with others with expertise. This knowledge will enable them to effectively navigate through the vast amounts of information and extract meaningful insights. Moreover, it is crucial to include relevant physical parameters, such as pH, temperature, salinity, and seasonality, as they can offer valuable context and further enrich the interpretation of the results. By incorporating these additional factors, researchers can gain a deeper understanding of the ecological dynamics at play and better comprehend the intricate relationships between different organisms and their environments. This holistic approach to data analysis not only broadens the scope of the study but also allows for more robust conclusions and potentially groundbreaking discoveries.

Overall, setting up an eDNA study requires careful planning and consideration of various factors. Defining the research question, selecting the appropriate sample type, implementing contamination reduction strategies, choosing suitable primers and databases, and conducting robust data analysis are all key considerations. By addressing these factors, one can ensure the success and reliability of their eDNA study and contribute valuable insights into the field of biodiversity monitoring and conservation.

15 | FUTURE ENDEAVORS

Given the success and interest of the first iteration of this workshop, future iterations of the Hong Kong International eDNA workshop are planned to be held biannually. We will be mindful of key changes in sequencing technologies and plan to incorporate more advanced in-field approaches to sequencing. We will also look to incorporate more direct experience with regard to experimental design and subsequent downstream analyses. As with this iteration, a pre-workshop survey will be utilized to gauge participant interest with regard to eDNA source material, organism(s) of interest, and sequencing methods to enable fine tuning of instruction and practical tailored to the interest of the participants.

With the world moving on from the events of COVID-19, it is paramount that in-person instruction be reintegrated into teaching practices in general. The overwhelming positive response of the participants to our hands-on instruction and the general social exchange held throughout the workshop was testament to the importance of human interaction in the learning and exchange of scientific insights and knowledge.

AUTHOR CONTRIBUTIONS

MS, IG, GJJ, MH, ML, VE designed the workshop and outlined the commentary. All authors contributed to writing and editing.

ACKNOWLEDGMENTS

We acknowledge The Croucher Foundation Limited (Hong Kong) for funding support via a Sponsorship of Conferences grant. Additional thanks to The eDNA Collaborative (Seattle, USA) for their input and discussions during the workshop. We also thank the members of the workshop committee, including David Baker, Shelby McIlroy, Kenneth Leung, Meng Yan, Jianwen Qiu, Jerome Hui, Jonathan Fong, Anthony Lau and Chi Chiu Cheang. We are also grateful to the eDNA and eEcology Lab members that helped with logistics including Kwan Yan Wong, Cheung Shuk Kwan, Ka Lam Karen Yeung John Ross Gabriel Sison Badajos, Si Zhewei, and Rachel Wong.

CONFLICT OF INTEREST STATEMENT

None.

DATA AVAILABILITY STATEMENT

There is no data associated with the commentary.

ORCID

Mathew Seymour https://orcid.org/0000-0002-3654-4857
Isis Guibert https://orcid.org/0000-0002-6896-6659
Gert-Jan Jeunen https://orcid.org/0000-0001-7458-3550
Micaela Hellström https://orcid.org/0000-0002-4704-724X
Matthieu Leray https://orcid.org/0000-0002-7327-1878
Chun Ming How https://orcid.org/0000-0001-5219-3826
Jack Chi Ho Ip https://orcid.org/0000-0002-4477-7351
Jessica Moffitt https://orcid.org/0000-0003-3825-2521
Xian Qin https://orcid.org/0000-0002-2109-8823

REFERENCES

Alberdi, A., & Gilbert, M. T. P. (2019). A guide to the application of Hill numbers to DNA-based diversity analyses. *Molecular Ecology Resources*, *19*(4), 804–817. https://doi.org/10.1111/1755-0998.13014

Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, *17*(1), 1–17.

Bohmann, K., Elbrecht, V., Carøe, C., Bista, I., Leese, F., Bunce, M., Yu, D. W., Seymour, M., Dumbrell, A. J., & Creer, S. (2022). Strategies for sample labelling and library preparation in DNA metabarcoding studies. *Molecular Ecology Resources*, *22*(4), 1231–1246. https://doi.org/10.1111/1755-0998.13512

Borcard, D., Gillet, F., & Legendre, P. (2018). *Numerical ecology with R* (Vol. *2*, 2nd ed.). Springer.

Bruce, K., Blackman, R., Bourlat, S. J., Hellström, A. M., Bakker, J., Bista, I., Bohmann, K., Bouchez, A., Brys, R., Clark, K., Elbrecht, V., Fazi, S., Fonseca, V., Hänfling, B., Leese, F., Mächler, E., Mahon, A. R., Meissner, K., Panksep, K., … Deiner, K. (2021). *A practical guide to DNA-based methods for biodiversity assessment*. Pensoft Publishers.

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D. M., de Vere, N., Pfrender, M. E., & Bernatchez, L. (2017). Environmental DNA metabarcoding:

Transforming how we survey animal and plant communities. *Molecular Ecology*, *26*(21), 5872–5895. https://doi.org/10.1111/mec.14350

Elbrecht, V., & Leese, F. (2017). PrimerMiner: An R package for development and in silico validation of DNA metabarcoding primers. *Methods in Ecology and Evolution*, *8*(5), 622–626.

Hakimzadeh, A., Abdala Asbun, A., Albanese, D., Bernard, M., Buchner, D., Callahan, B., Caporaso, J. G., Curd, E., Djemiel, C., Brandström Durling, M., Elbrecht, V., Gold, Z., Gweon, H. S., Hajibabaei, M., Hildebrand, F., Mikryukov, V., Normandeau, E., Özkurt, E., Palmer, M. J., … Anslan, S. (2023). A pile of pipelines: An overview of the bioinformatics software for metabarcoding data analyses. *Molecular Ecology Resources*.

Hleap, J. S., Littlefair, J. E., Steinke, D., Hebert, P. D., & Cristescu, M. E. (2021). Assessment of current taxonomic assignment strategies for metabarcoding eukaryotes. *Molecular Ecology Resources*, *21*(7), 2190–2203.

Jeunen, G. J., Dowle, E., Edgecombe, J., von Ammon, U., Gemmell, N. J., & Cross, H. (2023). CRABS—A software program to generate curated reference databases for metabarcoding sequencing data. *Molecular Ecology Resources*, *23*(3), 725–738.

Jeunen, G. J., von Ammon, U., Cross, H., Ferreira, S., Lamare, M., Day, R., Treece, J., Pochon, X., Zaiko, A., Gemmell, N. J., & Stanton, J.-A. L. (2022). Moving environmental DNA (eDNA) technologies from benchtop to the field using passive sampling and PDQeX extraction. *Environmental DNA*, *4*(6), 1420–1433.

Kvist, S. (2013). Barcoding in the dark?: A critical view of the sufficiency of zoological DNA barcoding databases and a plea for broader integration of taxonomic knowledge. *Molecular Phylogenetics and Evolution*, *69*(1), 39–45. https://doi.org/10.1016/j.ympev.2013.05.012

Leese, F., Sander, M., Buchner, D., Elbrecht, V., Haase, P., & Zizka, V. M. A. (2021). Improved freshwater macroinvertebrate detection from environmental DNA through minimized nontarget amplification. *Environmental DNA*, *3*(1), 261–276. https://doi.org/10.1002/edn3.177

Leray, M., Knowlton, N., & Machida, R. J. (2022). MIDORI2: A collection of quality controlled, preformatted, and regularly updated reference databases for taxonomic assignment of eukaryotic mitochondrial sequences. *Environmental DNA*, *4*(4), 894–907.

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., Boehm, J. T., & Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, *10*(1), 34. https://doi.org/10.1186/1742-9994-10-34

Lin, B., Dietrich, M. L., Senior, R. A., & Wilcove, D. S. (2021). A better classification of wet markets is key to safeguarding human health and biodiversity. *The Lancet Planetary Health*, *5*(6), e386–e394.

Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J., & Knight, R. (2011). UniFrac: An effective distance metric for microbial community comparison. *The ISME Journal*, *5*(2), 169–172. https://doi.org/10.1038/ismej.2010.133

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, *17*(1), 10–12.

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

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, *4*, e2584.

Seymour, M. (2019). Rapid progression and future of environmental DNA research. *Communications Biology*, *2*(1), 80. https://doi.org/10.1038/s42003-019-0330-9

Seymour, M., Edwards, F. K., Cosby, B. J., Bista, I., Scarlett, P. M., Brailsford, F. L., Glanville, H. C., de Bruyn, M., Carvalho, G. R., & Creer, S. (2021). Environmental DNA provides higher resolution assessment of riverine biodiversity and ecosystem function via spatio-temporal nestedness and turnover partitioning. *Communications Biology*, *4*(1), 512. https://doi.org/10.1038/s42003-021-02031-2

Seymour, M., & Smith, A. (2023). Arctic char occurrence and abundance using environmental DNA. *Freshwater Biology*, *68*(5), 781–789. https://doi.org/10.1111/fwb.14063

Spens, J., Evans, A. R., Halfmaerten, D., Knudsen, S. W., Sengupta, M. E., Mak, S. S. T., Sigsgaard, E. E., & Hellström, M. (2017). Comparison of capture and storage methods for aqueous macrobial eDNA using an optimized extraction protocol: Advantage of enclosed filter. *Methods in Ecology and Evolution*, *8*(5), 635–645. https://doi.org/10.1111/2041-210X.12683

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

Wang, H., Peng, Y., Wang, C., Wen, Q., Xu, J., Hu, Z., Jia, X., Zhao, X., Lian, W., Temmerman, S., Wolf, J., & Bouma, T. (2021). Mangrove loss and gain in a densely populated urban estuary: Lessons from the Guangdong-Hong Kong-Macao Greater Bay Area. *Frontiers in Marine Science*, *8*, 693450.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag.

Yates, M. C., Fraser, D. J., & Derry, A. M. (2019). Meta-analysis supports further refinement of eDNA for monitoring aquatic species-specific abundance in nature. *Environmental DNA*, *1*(1), 5–13.

How to cite this article: Seymour, M., Guibert, I., Jeunen, G.-J., Hellström, M., Leray, M., Elbrect, V., Lau, A., Cutracci, M., How, C. M., Ip, J. C. H., Liu, Z., Moffitt, J., Qian, C., Qin, X., Tse, C. M., Wang, Y., Wang, Z., Zeng, X., & Zhao, M. (2024). The First International eDNA Workshop in Hong Kong: A beginner's guide for the next-generation eDNA researcher. *Environmental DNA*, *6*, e552. https://doi.org/10.1002/edn3.552