Environmental DNA (eDNA) metabarcoding approach in fisheries research in India

*P. Jayasankar, M. A. Pradeep, K. G. Mini and T. V. Arunkumar
ICAR-Central Marine Fisheries Research Institute, Kochi
**e-mail : p.jayasankar@icar.gov.in*

Environmental DNA (eDNA) is defined as the genetic material obtained from a water sample containing no distinguishing signs of source macro-organisms. The method utilizes DNA which is continuously excreted by organisms into the surrounding environment through mucus, gamates, faeces, blood and other cells, and captures, analyses and obtains the nucleotide sequence of this DNA based on an environmental sample. eDNA analysis has emerged as a potentially powerful tool to access aquatic community structures. Analysis of eDNA can give us information on the organisms, their abundance and biomass through two approaches – eDNA barcoding and eDNA metabarcoding. In the former, specific species are targeted in samples using standard or quantitative PCR, and using traditional Sanger sequencing method. In the latter, the whole community is screened using multiple conserved primers and Next Gen Sequencing (NGS) is done. Studies suggest that eDNA metabarcoding outperforms traditional survey methods in terms of non-invasive sampling, sensitivity and cost incurred. There is now increased interest in using eDNA to supplement existing survey methods.

Since 2012 there has been a plethora of studies on eDNA metabarcoding as applied in biodiversity conservation, fish community identification, fisheries management, invasive species, as well as in fish biomass/abundance estimations (Hansen *et al.*, 2018 *Fish and Fisheries*, 1-18). A total 25 research papers related to eDNA metabarcoding/ metagenomics by Indian authors are predominantly pertaining to the study of microbial biodiversity from food, soil and deep sea sediments (Jiang and Yang, 2017 *Current Science* 112(8): 1659-1664). Not a single publication related to such study in fish has been cited from India.

Metabarcoding is constrained by factors like PCR efficiency, primer tags and sequencing efficacy. Another limitation is lack of comprehensively cured reference databases for certain metazoans for assigning taxon to the Operational Taxonomic Units (OTUS). Further studies are needed to improve sampling strategies (selection of season, sampling location within habitat, etc.) and to understand the relationship between sequence reads and species density. Gaps in knowledge about the dynamic mechanisms relating to shedding of tissue into the environment and metabolism related processes which could also affect quantity of DNA released by an organism into the water have to be filled. Dynamics of eDNA under field conditions, such as patterns of release, degradation, and diffusion will have to be taken into consideration to get estimates of fish distribution and biomass/abundance based on eDNA.

Methodology includes seawater filtration, quantitative real-time PCR, library preparation, Next Gen Sequencing (NGS) and statistical analysis. Copy number of DNA could be quantitatively interpreted in terms of fish abundance and biomass. High throughput sequencing data analysis using the state-of-the art tools could throw light on family level abundance in general and species level abundance of fish in particular. However, the strength of the relationship depends on environmental parameters, such as water temperature, and technical parameters, such as the filter being used for capturing eDNA. Species biology, environment and filtration methods and other

factors (e.g. extraction and fish ecology and spatial distribution) are likely to interact and significantly influence eDNA concentration variation. Caution is needed when interpreting the patterns of eDNA concentration in practical contexts. Parameters such as detection limits in water samples, influence of microbial activities on eDNA degradation, sampling design, seasonal conditions, nature of eDNA and fish ecology should be considered in future studies before predicting fish abundance from eDNA in natural conditions.

A basic study design and sampling strategies are essential for estimation of biomass using eDNA surveys. The decision on sample number and density within various habitats is an important aspect whiling developing a statistical sampling strategy. Further, the relationship between fish density and eDNA abundance depends on factors such as taxon- and age-specific shedding rates, specific eDNA degradation rates in the given environment and non-local eDNA transported with sea currents. The effect of these factors has to be measured and taken into account while analyzing the data.

A preliminary study was initiated on eDNA barcoding from known fish samples in marine aquarium tanks in the institute. Five hundred milliliters (ml) of water (salinity 35 ppt; water temperature 26.2^{o}C) were collected from each of 4 different tanks (Table 1). The pooled water sample was filtered through 0.45µ filter (Millipore), DNA extracted and PCR-amplified. DNA was purified and cloned, and initially 5 positive clones were sequenced.

Of the five clones sequenced, as per BLAST (NCBI) search, four of them belonged to silver moony (*Monodactylus argenteus*) and one to skunk clown fish (*Amphiprion akallopisos*). The results were in agreement with the dominance of species in the samples as given in Table 1, thereby confirming efficacy of the present methodology used. Although the scale of the aquarium experiment was much smaller than that of a natural population, the results provide basic information on the relationship between eDNA concentration and biomass. In addition, biotic and abiotic factors such as microbial activity, temperature, salinity and pH can influence eDNA survival and availability. Statistically, the effect of these factors on eDNA concentrations can be evaluated by using a general linear model. In general, type II regression model which can treat two variables with equal magnitude of random variation is be used to study the relationship between eDNA concentration and biomass of each species per milliliter water sample. After optimizing the methods to evaluate the concentration in a controlled aquarium, they can be tested in a natural environment. Research in this novel area can generate eDNA signatures of exploited fish species, which would facilitate for accurate estimation of biomass/abundance of fish in Indian seas. Further, India-specific eDNA-linked database on exploited marine species can also be developed.

There has been amazing advancement in technology from quantitative real time PCR to smart phone-powered sequencer, which would minimize many of the classical logistical and practical challenges of handling, storing and transport of

Table 1. Particulars of fish species sampled in the preliminary study

Tank #	Fish species	Number	Total length (mm)
1	Blackbar Triggerfish (*Rhinecanthus aculeatus*),	1	70
	Threespot Dascyllus (*Dascyllus trimaculatus*)	1	90
2	Yellow tail Angelfish (*Apolemicthys xanthurus*)	1	120
	Blue streak cleaner wrasse (*Labroides dimidiatus*)	1	80
	Canarytop wrasse *(Halichoeres leucoxanthus)*	1	90
3	Skunk Clown fish (*Amphiprion akallopisos*)	10	70-80
	Cerulean damsel (*Pomacentrus caeruleus*)	1	70
	Threespot Dascyllus (*Dascyllus trimaculatus*)	1	90
4	Silver moony (*Monodactylus argenteus*)	20	90-100

environmental samples. Finally technological advancement has reached in automated real time DNA measurements as in Environmental Sample Processor (ESP), which is set to monitor a specific geographic location ranging from coastal to deep sea, and does everything right from regular water sampling and storing to real-time molecular analysis. ESP may be costly, but cost-competitive compared to extensive ship time for visual monitoring or to continuous collection of water samples. eDNA is under the influence of many physical, chemical and biological parameters, which need to be analysed. Its role in direct quantitative assessment is still challenging. Current focus of research in this field should be around relative strengths on detection of presence/absence, migration patterns and life history events, broad ecological understanding, taxonomic coverage and providing basis for ecosystem-based management. Despite the caveats, eDNA-based monitoring will continue to develop to have profound impact on futuristic fisheries research and management.