Metabarcoding & Metagenomics

Case Study

The use of citizen science in fish eDNA metabarcoding for evaluating regional biodiversity in a coastal marine region: A pilot study

Masaki Miya[1], **Tetsuya Sado**[1], **Shin-ichiro Oka**[2], **Takehiko Fukuchi**[1]

1 *Natural History Museum and Institute, Chiba, Chiba 260-8682, Japan*

2 *Okinawa Churashima Foundation, Okinawa 905-0206, Japan*

Corresponding author: Masaki Miya (masaki_miya@me.com)

Abstract

To test the feasibility of a citizen science program for fish eDNA metabarcoding in coastal marine environments, we recruited six groups of voluntary citizens for a science education course at a natural history museum. We held a seminar on eDNA and a workshop for seawater sampling and on-site filtration using syringes and filter cartridges for the participants. After that, they selected single survey sites following the guidelines for conducting a safe field trip. They performed seawater sampling and on-site filtration at these sites during their summer holidays. The six selected sites unexpectedly included diverse coastal habitats within a 40 km radius, located at temperate latitudes in central Japan (~35°N). After the field trips, they returned filtered cartridges to the museum, and we extracted eDNA from the filters. We performed fish eDNA metabarcoding, along with data analysis. Consequently, we identified 140 fish species across 66 families and 118 genera from the six samples, with species richness ranging from 14 to 66. Despite its limited sample size, such a diverse taxonomic range of fish species exhibited spatial biodiversity patterns within the region, which are consistent with species distribution. These include north-south and urbanization gradients of species richness, geographic structure of the fish communities, and varying salinity preferences of the component species. This case study demonstrates the potential of fish eDNA metabarcoding as an educational and scientific tool to raise public awareness and perform large-scale citizen science initiatives encompassing regional, national, or global fauna.

Key Words

citizen science, conservation, eDNA metabarcoding, education, fish biodiversity, outreach

Introduction

Coastal marine regions constitute an interface between the vast open oceans and landmasses, encompassing various ecosystems – including rocky reefs, sandy shores, coral reefs, mangroves, seagrass beds, and estuaries (Isla 2009). We derive enormous benefits directly or indirectly from marine ecosystems, such as food supply, microclimate stabilization, and a cultural venue for recreational activities (Johns et al. 2014), which are called "ecosystem services" (Steiner 2014). The sustainable utilization of these ecosystem services necessitates the continuous monitoring of the status of

the marine environment, including abiotic and biotic components (Miya 2022). However, achieving this in the case of the latter components is difficult because they cannot be automatically monitored in the same way as the abiotic attributes (e.g., temperature and salinity).

In marine fishes, more than 16,700 species are known worldwide, with an average of approximately 100–150 new species being described annually (Eschmeyer et al. 2010). Traditionally, to investigate the species number and composition in specific marine areas with highly diversified fishes, field surveys that utilize direct capture-based sampling methods (e.g., netting and fishing) or underwater visual censuses (e.g., diving and underwater

PENSOFT.

video) are conducted (Oka et al. 2021). However, such field surveys are invasive, labor-intensive, time-consuming, costly, and require highly specialized taxonomic expertise for species identification (Thomsen and Willerslev 2015). Therefore, it is practically impossible to continuously monitor fish biodiversity using traditional approaches (Miya 2022).

Environmental DNA (eDNA) – defined as extra-organismal DNA left behind by macroorganisms (Bohmann et al. 2014) – has received increasing attention as an indirect genetic marker for inferring species presence for biodiversity monitoring (Deiner et al. 2017; Cristescu and Hebert 2018). In particular, the eDNA metabarcoding approach enables simultaneous detection of multiple species using a high-throughput sequencing platform (Taberlet et al. 2012). In fish, universal PCR primers amplifying a short hypervariable region of the mitochondrial DNA have been developed for species detection (Miya et al. 2015; Taberlet et al. 2018). The use of these primers in eDNA metabarcoding has been proven useful for addressing various ecological questions (Miya 2022).

The simplicity of the protocol used to collect eDNA samples from aquatic environments (Deiner et al. 2017), coupled with the eDNA metabarcoding approach, makes it possible to perform fish biodiversity monitoring in large spatio-temporal scales (Miya et al. 2020). Analysis of the resulting large datasets can reveal the spatio-temporal dynamics of fish communities, providing useful information for sustainable ecosystem services (Miya 2022). However, the potential benefits of eDNA metabarcoding for biodiversity monitoring are not well recognized (Deiner et al. 2017). Thus, public awareness of the potential of eDNA metabarcoding must be raised and effective government policies for the same must be proposed (Hupało et al. 2021; Yang et al. 2021). In addition, the simplicity of the protocol for collecting eDNA samples can provide a good opportunity for developing citizen science programs (Deiner et al. 2017). The involvement of citizen scientists can consequently reduce the financial cost of biodiversity monitoring, while increasing stakeholder engagement and approval ("Reef Vision"; Florisson et al. 2018). Indeed, the eDNA Society of Japan has been conducting fish biodiversity surveys using fish eDNA metabarcoding at the national level in collaboration with the National Museum of Emerging Science and Innovation, and the results are published on the web site (https://ednasociety.org/#news-letter).

A major purpose of this study was to test the feasibility of the eDNA metabarcoding approach as a biodiversity monitoring tool in a citizen science initiative at the regional level using a regular science education course at a public natural history museum. Due to the COVID-19 pandemic, we recruited only six groups of parents and children. These six groups collected and filtered seawater samples from diverse coastal habitats within a 40 km radius during their summer holidays. The fish eDNA metabarcoding of the six samples detected a diverse taxonomic range of fish, including 140 species across 66 families and 118 genera. Subsequent ecological analyses revealed distinct geographic structures of species richness and fish communities, demonstrating the potential of fish eDNA metabarcoding as an educational and citizen science tool.

Methods

Ethics statement

All seawater sampling at the six sites was conducted in compliance with Japanese laws and regulations as well as local ones.

Collection of eDNA samples

As part of the various outreach activities at the Natural History Museum and Institute, Chiba, the present study recruited six pairs of parents and children on the website (http://www2.chiba-muse.or.jp/NATURAL/) to collect eDNA samples from marine environments in Chiba and the surrounding prefectures. A seminar and a workshop were conducted on July 4, 2021, to provide an overview on eDNA studies and to secure high-quality eDNA samples. A sampling kit was distributed to each group during the workshop and its use was demonstrated. Moreover, examples of appropriate or inappropriate sites for water sampling in coastal areas were demonstrated, but no specific regions or locations were mentioned to the participants. Thus, they freely chose water sampling sites, depending on their personal preferences of residence and summer holiday locales. All relevant collection data are summarized in Suppl. material 1: Table S1.

For seawater collection, low-tech bucket sampling was utilized, comprising various small commercial buckets fastened to a rope (provided by the participants). Prior to seawater sampling, the participants wore disposable gloves on both hands and assembled a set of on-site filtration kits consisting of a Sterivex filter cartridge (pore size 0.45 μm; Merck Millipore, MA, USA) and a 50-ml disposable syringe with a Luer lock connector (TERUMO, Tokyo, Japan). Subsequently, they fixed the end of the rope fastened to the bucket and collected surface seawater by casting and retrieving the bucket filled with seawater.

The participants performed on-site filtration using the above kit (filter cartridge + syringe) to collect and concentrate the eDNA on the filter membrane inside the cartridge. They removed the filter cartridge from the syringe and drew approximately 50-ml of seawater into the syringe by pulling the plunger, reattaching the filter cartridge to the syringe, and pushing the plunger for filtration of the seawater. This step was repeated until the final filtration volume reached 1000 mL. When the filter was clogged before reaching 1000 mL filtration, the total volume of filtered seawater was recorded.

After on-site filtration, an outlet port of the filter cartridge was closed with a 3-mm diameter rubber cap

(KOKUGO, Tokyo, Japan), filled with 1.6-mL RNAlater (Thermo Fisher Scientific, DE, USA) into the cartridge from an inlet port of the cartridge using a disposable capillary pipette (AS ONE, Tokyo, Japan) to prevent eDNA degradation, and an inlet port was closed with a screw cap (TERUMO, Tokyo, Japan) for preservation. They transported the filtered cartridges to their homes in a portable cooler with ice packs and kept them at 4 °C in their fridges prior to shipment. Subsequently, they returned the cartridges to the museum using a refrigerated courier, and the six cartridges were stored at –20 °C until eDNA extraction was performed.

For more details of the eDNA collection, see "Environmental DNA sampling and experiment manual Version 2.1 (The eDNA Society 2019)

Laboratory protocol

Details of the laboratory protocol can be found in the Suppl. materials 5: Supplementary methods. The workspace and equipment were thoroughly sterilized prior to all laboratory experiments. Filtered pipette tips were used and eDNA-extraction, pre- and post-PCR manipulations were conducted in three different dedicated rooms that were physically separated from each other to safeguard against cross-contamination from PCR products.

eDNA was extracted from the filter cartridges using a DNeasy Blood & Tissue kit (Qiagen, Hilden, Germany) following the methods developed by Miya et al. (2016) with slight modifications. After removing the preservative in the cartridge, the filter cartridge was lysed using proteinase K. The collected DNA extract (~900 μL) was purified using the DNeasy Blood and Tissue kit following the manufacturer's protocol and the final elution volume was set to 200 μL. An extraction blank (EB) was also created during this process.

This study employed two-step PCR for paired-end library preparation using the MiSeq platform (Illumina, CA, USA). In general, we followed the methods developed by Miya et al. (2015) and subsequently modified by Miya and Sado (2019). For the first round of PCR (1st PCR), a mixture of the following six primers was used: MiFish-U-forward, reverse, MiFish-E-forward-v2, reverse-v2, MiFish-U2-forward, and reverse. These primer pairs amplify a hypervariable region of the mitochondrial 12S rRNA gene (ca. 172 bp; hereafter called "MiFish sequence") and append primer-binding sites (5´ ends of the sequences before six Ns) for sequencing at both ends of the amplicon.

The 1st PCR was performed with eight technical replicates for the same eDNA template to minimize PCR dropouts. A 1st PCR blank (1B) was also prepared during this process, in addition to EB. After completing the 1st PCR, an equal volume of PCR products was pooled from each of the eight replicates, and the pooled products were purified, quantified, and diluted to 0.1 ng/μL using Milli Q water, and the diluted products were used as templates for the second round of PCR (2nd PCR). For the two blanks (EB, 1B), the 1st PCR products were purified in the same manner.

The 2nd PCR was performed to append dual-index sequences and flow cell binding sites for the MiSeq platform. A 2nd PCR blank (2B) was also prepared during this process, in addition to EB and 1B. The six PCR products as well as the three blank samples from the 2nd PCR products were pooled along with other samples from different projects. Subsequently, the pooled dual-indexed libraries were electrophoresed on an agarose gel, and the target amplicons (~370 bp) were excised. The concentration of the size-selected libraries was measured, diluted to 10.0 pM, and sequenced on the MiSeq platform using a MiSeq v2 Reagent Kit for 2 × 150 bp PE (Illumina, CA, USA) following the manufacturer's protocol.

All raw DNA sequence data and associated information were deposited in the DDBJ/EMBL/GenBank database and are available under accession number DRA012840.

Sequence analysis

Data preprocessing and analysis of raw MiSeq reads from the MiSeq run were performed using PMiFish ver. 2.4 (https://github.com/rogotoh/PMiFish.git; Miya et al. 2020). Forward (R1) and reverse (R2) reads were merged while discarding low-quality tail reads with a cut-off threshold set at a quality (Phred) score of 2, too short reads (<100 bp) after tail trimming, and those paired reads with too many differences (>5 positions) in the aligned region (ca. 65 bp). Primer sequences were removed from those merged reads and those reads without the primer sequences underwent quality filtering to remove low quality reads. The preprocessed reads were dereplicated, and all singletons, doubletons and tripletons were removed from the subsequent analyses to avoid false positives (Edgar 2010). The dereplicated reads were denoised to generate amplicon sequence variants (ASVs) that removed all putatively chimeric and erroneous sequences (Callahan et al. 2017).

The ASVs were subjected to taxon assignments to species names (molecular operational taxonomic units; MOTUs) with a sequence identity of >98.5% with the reference sequences (two nucleotide differences allowed) and a query coverage of ≥90%. An incomplete reference database necessitates this clustering step, which enables the detection of multiple MOTUs for identical species names. Such multiple MOTUs were annotated with "gotu1, 2, 3…" and all of these outputs (MOTUs plus U98.5 MOTUs) were tabulated with read abundances. ASVs with sequence identities of <80% (saved as "no hit") were excluded from the above taxon assignments and downstream analyses because all of them were found to be non-fish organisms. MiFish DB ver. 43 was used for taxon assignment, comprising 7973 species distributed across 464 families and 2675 genera.

To refine the above taxon assignments, family level phylogenies were reproduced from MiFish sequences from MOTUs and reference sequences (contained in the MiFish DB ver. 43) belonging to these families. For each family, representative sequences (most abundant reads)

from MOTUs were assembled, all reference sequences were added from that family, and saved in FASTA format. The combined FASTA-formatted sequences were subjected to multiple alignments using MAFFT 7 (Katoh and Standley 2013) with a default set of parameters. A neighbor-joining (NJ) tree was subsequently constructed with the aligned sequences in MEGA X (Stecher et al. 2020) using Kimura two-parameter distances.

A total of 67 family-level trees were visually inspected and taxon assignments were revised in the following manner. For those U98.5 MOTUs placed within a monophyletic group consisting of a single genus, the unidentified MOTUs were named after that genus, followed by "sp." with sequential numbers (e.g., *Pagrus* sp. 1, sp. 2, sp. 3...). For the remaining MOTUs ambiguously placed in the family-level tree, the unidentified MOTUs were named after that family, followed by "sp." with sequential numbers (e.g., Sparidae sp. 1, sp. 2, sp. 3...).

The final list of detected species can be found in Suppl. material 2: Table S2.

Fish community analysis

The differences in fish community structures were visualized using non-metric multidimensional scaling (NMDS) with 999 separate runs of real data. For NMDS, community dissimilarity was calculated based on incidence-based Jaccard indices, and NMDS stress was used to confirm the representation of NMDS ordination. The "metaMDS" function of "vegan" ver. 2.5.6 package (Oksanen et al. 2019) was used for NMDS.

To evaluate whether the detected fish communities reflected the marine environments of the six sampling sites (Fig. 1), three types of salinity preferences (salt, salt/brackish, salt/brackish/freshwater) were assigned for each detected species using information available from FishBase (Froese and Pauly 2019). A PMiFish pipeline (ver. 2.4) automatically retrieved the relevant information (water area, salinity preference, depth range) through FishBase in a default setting using "rfishbase" (Boettiger et al. 2012).

To evaluate the biogeographic characteristics of the detected fish communities from the six sampling sites, the center of the geographic distribution for each detected species was calculated by averaging the latitudes of the northern and southern limits in the Northern Hemisphere following Masuda (2008), with reference to the distribution data in Nakabo (2013). Those data can be found in Suppl. material 3: Table S3.

As a baseline dataset, a faunal inventory was compiled for coastal marine fish from the region including the survey areas based on multiple museum collections and literature surveys (Suppl. material 4: Table S4). Five collections were selected from neighboring natural history museums to represent coastal marine fish fauna in Chiba Prefecture: Natural History Museum and Institute, Chiba (CBM) and its Coastal Branch (CNMH), National Museum of Nature and Science, Tokyo (NSMT), Kanagawa Prefectural Museum of Natural History (KPM), and Yokosuka City Museum (YCM). Subsequently, those fishes collected from the coastal marine areas of Chiba prefecture were extracted from databases of these collections using the search word "Chiba," wherein fishes inhabiting freshwaters only were manually excluded (Suppl. material 4: Table S4). For the literature surveys, eight core references were designated (Miya et al. 1994a, b, 1995; Hagiwara and Kimura 2006; Kohno et al. 2011; Nakabo 2013; Aoki et al. 2016; Shitamitsu et al. 2019).
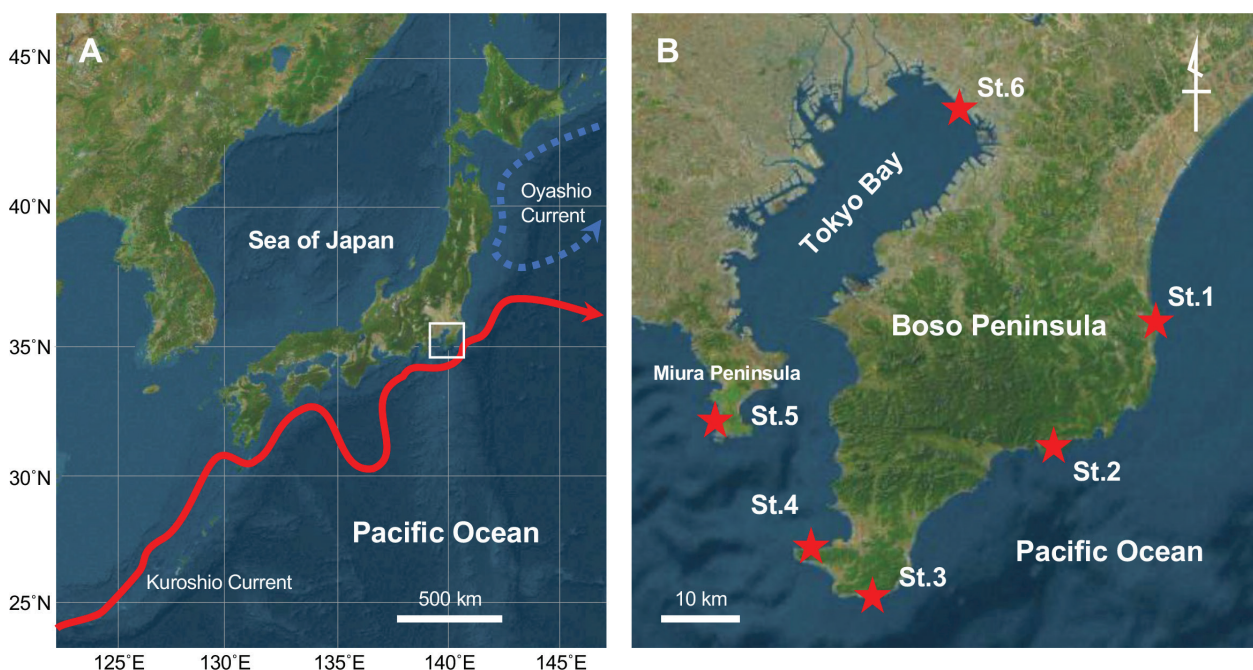
Figure 1. A) the location of the survey area and schematic flow paths of the Kuroshio (red) and Oyashio (blue) currents; **B**) the location of the six survey sites (Sts. 1–6) in Boso and Miura peninsulas. Map data 2021 Apple Inc.

Results

Collection of eDNA samples

Six groups of parents and children were selected from 22 groups of applicants, such that their places of residence were scattered. Those six groups freely chose their sampling sites (Sts. 1–6 in Fig. 1), performing seawater sampling and on-site filtration during the summer holiday of 2021, from July 26 to August 21 (Suppl. material 1: Table S1).

St. 1 is located at the southern end of Kujukuri Beach – a shallow, vast sandy beach with a total length of 66 km along the Pacific coast – and is strongly affected by the intrusion of the cold Oyashio water along the coastline (Yang et al. 1993). Although the seabed of St. 1 is entirely sandy, wave-dissipating blocks protect the shore from coastal erosion (Fig. 2). Moreover, the seawater is often turbid due to the inflow of muddy water from an estuary of the Isumi River approximately 1 km south (Fig. 2). The

filtered cartridge appeared darkest (Fig. 2) and clogged at a filtration volume of 600 mL

Sts. 2–5 are located in the southern tips of the Boso and Miura peninsulas, and are greatly affected by the warm Kuroshio current flowing northward along the Pacific coast as well as its branches (Soh 2003; Fig. 1). In contrast to St. 1, these sites are located within or surrounded by rocky shores with intricate coastlines (Fig. 2), and the seawaters are generally clear, even within fishing ports (Sts. 2 and 4). The filtered cartridges remained relatively clear and were not clogged, with a maximum filtration volume of 1000 mL (Fig. 2).

St. 6 is located at the mouth of the Hanami River, deep within Tokyo Bay (Figs 1, 2). The innermost part of Tokyo Bay was originally a vast tidal flat, and the natural coast has disappeared because of repeated reclamation and revetment for over 350 years (Endoh 2004), transforming into a large industrial area in the last 50 years (Fig. 2). The filtered cartridge appeared darker, comparable to St. 1 (Fig. 2), and became clogged at a filtration volume of 600 mL.
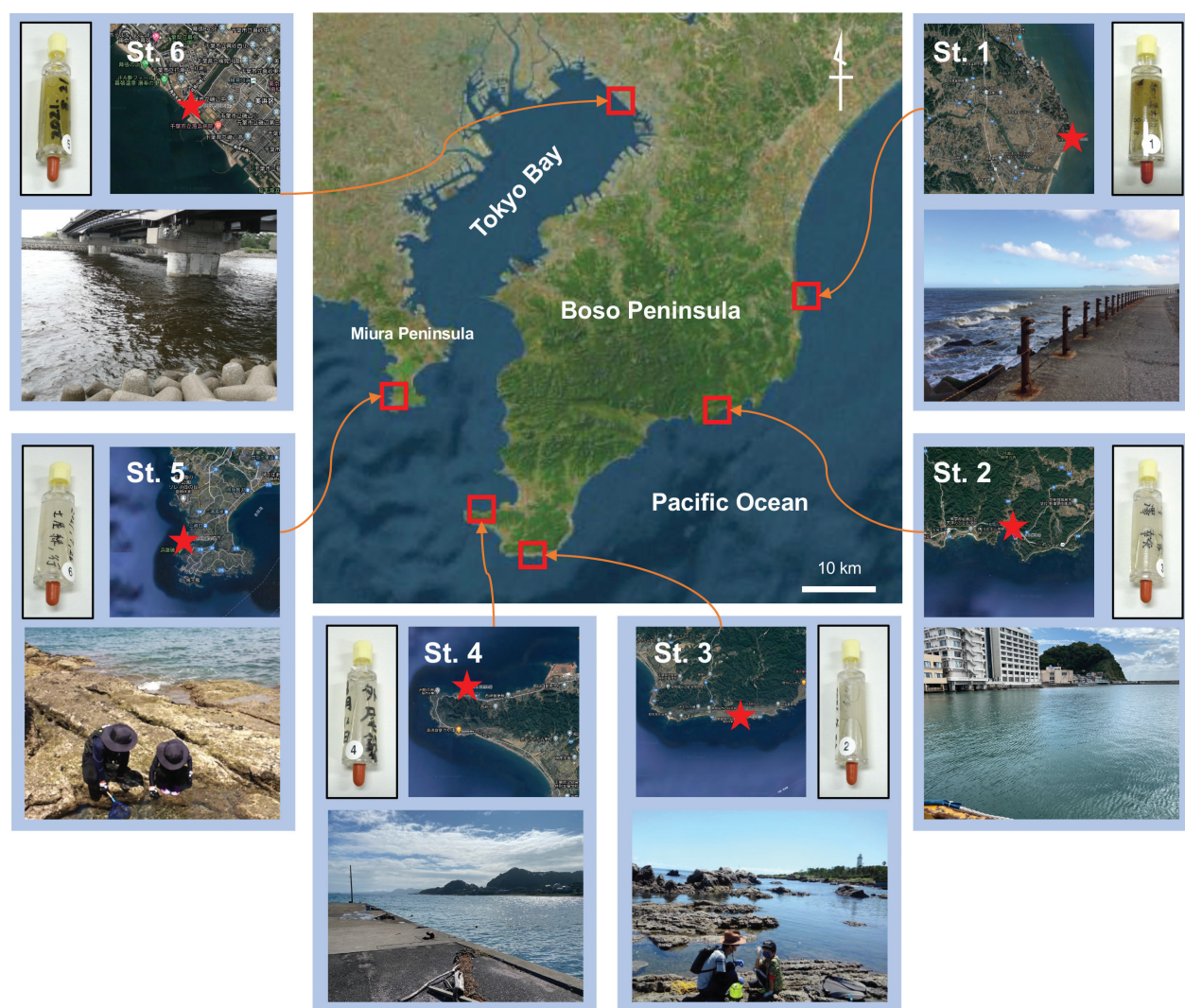
Figure 2. Enlarged maps of the six survey sites (Sts. 1–6), their snapshots, and corresponding filtered cartridges. The participants took photos from Sts. 1, 3, 5, and 6 on the survey date (Suppl. material 2: Table S2), and those of Sts. 2 and 4 were taken by one of us (MM) on September 28, 2021. Map data 2021 Apple Inc.

Library preparation and MiSeq sequencing

Based on the eDNA extracts, MiFish fragments were successfully amplified in the 1st PCR, with distinct, putatively fish bands being observed around ca. 310 bp (Miya et al. 2020) for all the six samples. Dual-index sequences and flowcell-binding sites were appended in the 2nd PCR and the target amplicons (ca. 370 bp) were purified and quantified for the six libraries. MiSeq paired-end sequencing (2 × 150 bp) of the six libraries – in conjunction with an additional 29 libraries (total = 35) – yielded a total of 6,135,029 reads, with an average of 97.2% base calls with Phred quality scores of ≥30.0 (Q30; error rate = 0.1% or base call accuracy = 99.9%). This run was highly successful, considering that the manufacturer's guidelines (Illumina Publication no. 770-2011-001 as of May 27, 2014) are >80% bases ≥Q30 at 2 × 150 bp.

Of the 6,135,092 reads, 531,306 were assigned to the six libraries, and the number of raw reads for each library ranged from 67,618 to 123,743, with an average of 88,551 reads. After merging two overlapping paired-end fastq files [654,497 reads (98.5%)], the primer-trimmed sequences were subjected to quality filtering to remove low-quality reads [650,461 reads (97.9%)]. The remaining reads were dereplicated for subsequent analysis and single to tripletons were removed from the unique sequences. Thereafter, the reads were denoised to remove putatively erroneous and chimeric sequences. The remaining 567,980 reads (85.5% of the raw reads) were subjected to taxon assignments after rarefaction to the minimum number of reads (67,681). Of these, 531,306 reads (93.5% of the denoised reads) were putatively considered fish sequences, while BLAST searches indicated that non-fish sequences [36,674 reads (6.5%)] primarily consisted of mammals (i.e., cows, pigs, and humans) and a few unknown sequences. The three negative controls (EB, 1 B, and 2 B) were subjected to the same analysis pipeline, whereby they did not yield any denoised reads.

Following the automatic taxon assignments, the family level NJ trees were visually inspected and the species names in the list were revised. The final list included 258 detections, assigned to 140 species across 66 families and 118 genera (Suppl. material 2: Table S2). All detected species can be found in the fish inventory comprising 948 species across 158 families and 493 genera (Suppl. material 4: Table S4), with the exception of unnamed species (family or generic names + sp.). The latter may be attributed to the lack of reference sequences or congeneric species with little or no interspecific sequence variations in the MiFish fragments, such as the rockfish genus *Sebastes* (Sebastidae), some flying fishes (Exocoetidae), the surfperch genus *Ditremma* (Embiotocidae), and the pufferfish genus *Takifugu* (Tetraodontidae) (Miya et al. 2020).

Fish community analysis

The number of species detected at the six sampling sites ranged from 14 (St. 1) to 66 (St. 5), with a mean of 43 (Suppl. material 2: Table S2). The number at the four sites along the rocky shores (49–66 spp. at Sts. 2–5) exceeded the two sites at a sandy shore (14 spp. at St. 1) and an inner bay (19 spp. at St. 6) (Fig. 3). For convenience, the rocky shores, sandy shore, and inner bay are hereafter called "area(s)" individually or collectively.

Of the 140 detected species, only Blackhead Seabream (*Acanthopagrus schlegelii*) and Grass Puffer (*Takifugu niphobles*; putatively including few congeneric species owing to low interspecific genetic sequence variations; Miya et al. 2020) were observed across the three areas. The former species is one of the most popular targets for anglers in Japan, while the latter is a small, ubiquitous poisonous fish that is not of much importance to anglers (https://fishing-fishing.com/2020/01/22/5783/). Approximately half of the detected species at the sandy-shore and inner-bay sites (7 of 14 spp. at St. 1 and 10 of 19 spp. at St. 6) were also identified in other areas, whereas 89% of the detected species at the rocky shore sites (110 of 123 spp. at St. 2–6) were unique to this area (Fig. 4). Within the latter area (rocky shore), 13–17 species (22–33%) were unique at each site, and the rest of the species were also found at two to four sites in this area. Furthermore, NMDS ordination analysis indicated that dissimilarities in the fish communities of these three areas were well separated in two-dimensional space (Fig. 5; NMDS stress = 0).

As expected from the geographical locations of the sampling sites (Fig. 2), fish tolerant of freshwater and salt-
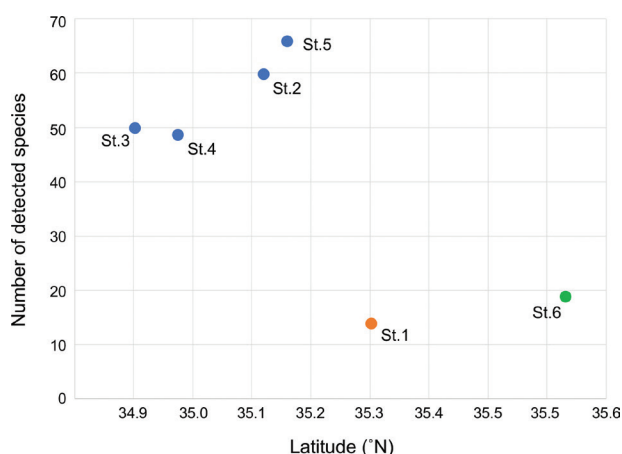
Figure 3. Scatter plot of the numbers of fish species detected from the fish eDNA metabarcoding. The numbers were not corrected for the filtered volumes (1000 mL for Sts. 2–5 and 600 mL for Sts. 1, 2) and are plotted against latitudes. Orange, blue, and green dots represent the data from sandy-shore, rocky-shore, and inner-bay areas, respectively. Note that the rocky-shore areas are mostly located along the southern coastline, while sandy-shore and inner-bay areas are located in the northern coastlines of Boso Peninsula. Thus, the gradient of the species richness does not necessarily relate to latitudes.
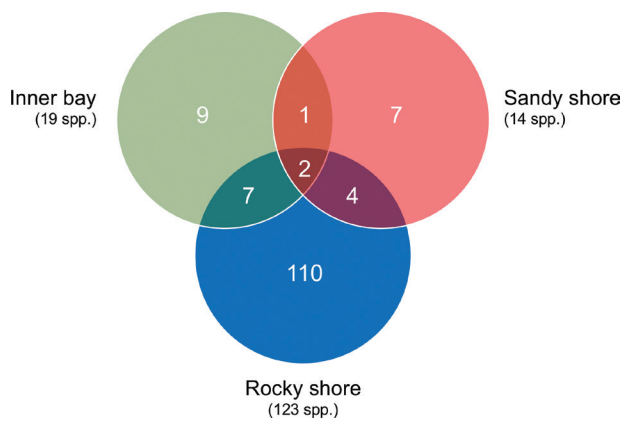
Figure 4. Venn diagram of the number of detected species from sandy-shore, rocky-shore, and inner-bay areas. Total numbers of the detected species for each area are indicated within parentheses.
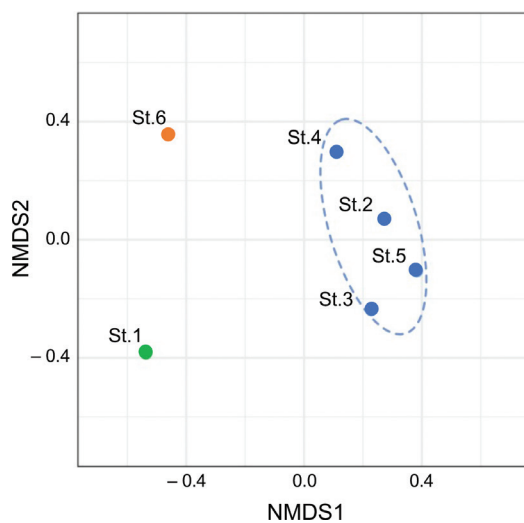
Figure 5. Non-metric multi-dimensional scaling (NMDS) for fish communities detected in four sites along rocky-shore (Sts. 2–5, blue), sandy-shore (St. 1, green) and inner-bay (St. 6, orange) areas. NMDS stress ≈ 0.

water accounted for approximately one-third of Sts.1 and 6 (Fig. 6). Of these fish species populating multiple habitats, Flathead Mullet (*Mugil cephalus cephalus*) and Yellowfin Goby (*Acanthogobius flavimanus*) appeared in common, although the marine environments of the two stations were significantly different (eroded sandy shore facing open ocean near the estuary vs. river mouth in the highly urbanized bay area). Other marine and brackish water species detected at the two stations (Suppl. material 2: Table S2) were frequently found in museum collections and references (Suppl. materials 2, 4: Tables S2, S4). In contrast to these two sites, marine fish were predominant at Sts. 2–5, occupying >90% of the component species (Fig. 3). Most of them were commonly found in rocky shores along the southern coast of Boso and Miura peninsulas (Miya et al. 1994a, b, Hagiwara and Kimura 2006; Shitamitsu et al. 2019) (Suppl. material 2, 4: Tables S2, S4).

The center of the geographic distribution of each component species was calculated from the average latitude of the northern and southern limits from literature records (Suppl. material 3: Table S3). The distribution patterns have been summarized in a box plot for each station (Fig. 7). The average latitudes of the component fish species from St. 1 – which are strongly affected by the cold Oyashio water – are biased to the north, while those from Sts. 2–5, which are strongly affected by the warm Kuroshio Current, were slightly biased to the south. Such a trend is ambiguous in the highly urbanized site (St.6), despite its northernmost location.

Discussion

The six selected sites unexpectedly included diverse coastal habitats within a 40 km radius – from sandy and rocky shores along the Pacific coast to a highly urbanized area deep within Tokyo Bay – located at temperate latitudes in central Japan (around 35°N; Fig. 1). Despite the limited number of sites and samples, eDNA metabarcoding analysis based on 567,980 reads detected 140 coastal fish species across 66 families and 118 genera from these libraries (Suppl. material 2: Table S2). To the best of our knowledge, the observed species richness (140 spp.) is the highest in coastal marine environments as revealed by fish eDNA metabarcoding, except for that of coral reefs (Jerde et al. 2019; Miya 2022). We acknowledge that greater survey effort (e.g., collecting more field samples of larger volume at each site) has been shown to increase the probability of detecting fish DNA, reducing the impact of false negatives and improving confidence in the eDNA metabarcoding approach (Ficetola et al. 2015; Pawlowski et al. 2018; Doi et al. 2019). However, we chose a single sample from each site to ensure safe field trips for the participants during the hot summer days.

This diverse taxonomic range of fish species exhibits distinct spatial biodiversity patterns within the region. For example, north-south and urbanization gradients of species richness are apparent (Fig. 3), which may reflect a complex mixture of habitat characteristics (sandy vs. rocky shores and natural vs. artificial coastal environments) (Fig. 4) and influences of cold waters (Oyashio) and warm currents (Kuroshio) along the coastal regions of the Boso Peninsula (Fig. 1). Indeed, habitat complexity – which generally increases fish biodiversity (Gratwicke and Speight 2005) – increases toward the south along the Pacific coast from sandy to rocky shores and in the Tokyo Bay area from artificial to natural coastal environments (Fig. 1). In addition, the warm Kuroshio Current transports tropical and subtropical species to the north (Senou et al. 2006; Saito 2019), resulting in clear north-south gradients in marine fish species richness (Tittensor et al. 2010). It should be noted that warm-water fish families, such as Labridae (including "Scaridae"; wrasses, eight spp.), Carangidae (jacks, seven spp.) and Kyphosidae (pilotfishes, six spp.) represent the three most speciose families that contributed to the 110 species
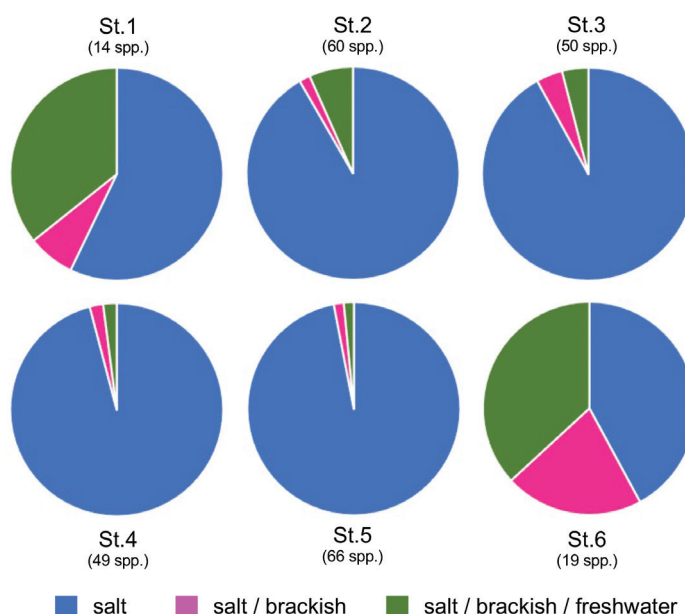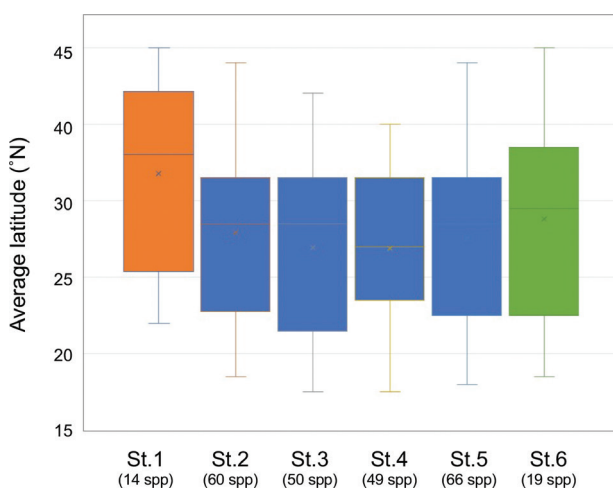
Figure 6. Proportions of the three categories of salinity preferences for detected fish species from each study site. Salinity-preference data were taken from FishBase (Froese and Pauly 2019) using "rfishbase" (Boettiger et al. 2012).

uniquely detected in the rocky-shore sites (Sts. 2–5, Fig. 4, Suppl. material 3: Table S3). The southern tip of the Boso Peninsula is also known to be the northernmost locality in the Pacific where coral communities are regularly observed (Yamano et al. 2012).

In addition to species richness, the composition of fish communities is geographically structured within this small region – those from rocky shores (Sts. 2–5, Fig. 4) were clustered and distantly located from the remaining two sites (Sts, 1, 6) in the two-dimensional space (Fig. 5). Furthermore, fish eDNA metabarcoding captures differences in salinity preferences of the component species, with those fishes using multiple habitats (from freshwater

to saltwater) being predominant at St. 1 (near the estuary) and St. 6 (at the river mouth), whereas marine fishes are predominant at Sts. 2–5 (Fig. 6). Of these fishes, the eDNA detection of chum salmon (*Oncorhynchus keta*) at St. 1 during the summer month (August) was the most noteworthy because it corresponded to the southern limit of this species' range just prior to the upstream migration period (Kajiyama 2009) when the species is difficult to observe or capture using the traditional methods. Patterns of the variations as seen in the center of the geographic distribution (as an average latitude of the northern and southern limits; Nakabo 2013) of each component species also reflected the characteristics of fish communities (Fig. 7), with more southern and northern species being predominant at Sts. 2–5 and St. 1, respectively, while it is ambiguous at the highly urbanized site in Tokyo Bay (St. 1). These observations on the regional characteristics of species diversity and fish communities (Figs 5–7) are concordant with previous records of fish fauna from these areas (Miya et al. 1994a, b, 1995; Hagiwara and Kimura 2006; Kohno et al. 2011).

Concluding remarks

Following the MiFish eDNA metabarcoding of the six samples and subsequent fish diversity and community analyses, a post-workshop event was conducted for the participants at our museum on September 19, 2021, to provide an overview of the results of the present surveys using PowerPoint slides. First, photos of the six filtered cartridges before eDNA extraction were shown and the differences in colors that reflect the turbidity of the seawater were compared (Fig. 2). The geographic and oceanographic backgrounds of Sts. 1 to 6 were introduced, along with the top 14 fish species based on

Figure 7. Box plots depicting variations in the center of geographic distribution for each detected species, calculated from an average of the latitudes of the southern and northern limits in the Northern hemisphere provided in Nakabo (2013). Orange, blue, and green bars represent the data from sandy-shore, rocky-shore, and inner-bay areas, respectively. Total numbers of the detected species for each area are indicated within parentheses.

read abundances, and topical issues on some of those fishes characteristic to the site from biological as well as cultural perspectives. We proceeded with this site-by-site presentation while asking the participants about their experiences during the surveys (Suppl. material 1: Table S1). In addition, the results of the fish diversity and community analyses, as shown in Figs 3–7, were demonstrated to summarize the significance of the entire survey. The workshop ended with a question-and-answer session, during which the participants asked a wide variety of questions concerning eDNA metabarcoding and fish diversity. This study did not intend to assess the educational effects of eDNA sampling and collection by the participants or subsequent demonstration of the results from eDNA metabarcoding analyses. Therefore, it is uncertain how these outreach activities have deepened the participants' understanding of the significance of fish biodiversity and conservation. However, the most impressive outcome was that the participants were uniformly amazed that eDNA metabarcoding could detect such a remarkable variety of fish species in just a cup of seawater.

This study demonstrated that the simplicity of the protocol for eDNA sampling allows citizen scientists (six groups of parents and children) to collect high-quality eDNA from coastal marine environments. Despite the limited sample size due to the ongoing COVID-19 pandemic, eDNA metabarcoding from the six samples detected a diverse taxonomic range of 140 fish species across 66 families and 118 genera. The number of fish species represents only a small fraction of the entire regional fauna (948 species across 158 families and 493 genera; Suppl. material 1: Table S1) compiled from museum collections and literature surveys. However, with this small eDNA metabarcoding dataset, we were able to show distinct spatial biodiversity patterns within the region – including north-south and urbanization gradients of the species richness (Figs 3, 4), geographic structure of the fish communities (Fig. 5), different salinity preferences of the component species (Fig. 6), and variations in the center of geographic distributions of the component species (Fig. 7), all of which are concordant with the previous knowledge of the regional fish fauna.

Thus, this case study shows the potential of fish eDNA metabarcoding as an educational tool to raise public awareness and deepen the understanding of fish biodiversity. This case study also highlights fish eDNA metabarcoding as a future scientific tool to perform large-scale citizen science initiatives covering the regional, national, or global fauna, if the sampling sites are properly selected, and citizen engagement is large enough (Meyer et al. 2021). In this regard, the role of local natural history museums in eDNA metabarcoding is critical because they deliver a wide range of field-based and online citizen science projects and play a central role in supporting the development of citizen science and citizen scientists (Sforzi et al. 2018). Finally, we should note that UNESCO recently started "eDNA Expedition" (https://ednaexpeditions.org), a new two-year project running from 2022 to 2023. This initiative aims at describing baseline coastal marine biodiversity using fish eDNA metabarcoding through local citizen engagement across UNESCO-listed 50 marine World Heritage sites.

Acknowledgements

References

Aoki T, Usui S, Kanai T, Aoki S, Okamoto K, Sano M (2016) Comparison of fish assemblage structures between exposed and protected sandy beaches at Uchibo, Boso Peninsula, central Japan. Nippon Suisan Gakkaishi 82(4): 569–580. https://doi.org/10.2331/suisan.16-00008

Boettiger C, Lang DT, Wainwright PC (2012) rfishbase: Exploring, manipulating and visualizing FishBase data from R. Journal of Fish Biology 81(6): 2030–2039. https://doi.org/10.1111/j.1095-8649.2012.03464.x

Bohmann K, Evans A, Gilbert MTP, Carvalho GR, Creer S, Knapp M, Yu DW, de Bruyn M (2014) Environmental DNA for wildlife biology and biodiversity monitoring. Trends in Ecology & Evolution 29(6): 358–367. https://doi.org/10.1016/j.tree.2014.04.003

Callahan BJ, McMurdie PJ, Holmes SP (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. The ISME Journal 11(12): 2639–2643. https://doi.org/10.1038/ismej.2017.119

Cristescu ME, Hebert PDN (2018) Uses and misuses of environmental DNA in biodiversity science and conservation. Annual Review of Ecology, Evolution, and Systematics 49(1): 209–230. https://doi.org/10.1146/annurev-ecolsys-110617-062306

Deiner K, Bik HM, Mächler E, Seymour M, Roussel AL, Altermatt F, Creer S, Bista I, Lodge DM, Vere N, Pfrender ME, 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

Doi H, Fukaya K, Oka S, Sato K, Kondoh M, Miya M (2019) Evaluation of detection probabilities at the water-filtering and initial PCR steps in environmental DNA metabarcoding using a multispecies site occupancy model. Scientific Reports 9(1): e3581. https://doi.org/10.1038/s41598-019-40233-1

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics (Oxford, England) 26(19): 2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Endoh T (2004) Historical review of reclamation works in the Tokyo Bay area. The Journal of Geography 113(6): 785–801. https://doi.org/10.5026/jgeography.113.6_785

Eschmeyer WN, Fricke R, Fong JD, Polack DA (2010) Marine fish diversity: History of knowledge and discovery (Pisces). Zootaxa 2525(1): 19–50. https://doi.org/10.11646/zootaxa.2525.1.2

Ficetola GF, Pansu J, Bonin A, Coissac E, Giguet-Covex C, Barba MD, Gielly L, Lopes CM, Boyer F, Pompanon F, Rayé G, Taberlet P (2015) Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. Molecular Ecology Resources 15(3): 543–556. https://doi.org/10.1111/1755-0998.12338

Florisson JH, Tweedley JR, Walker THE, Chaplin JA (2018) Reef vision: A citizen science program for monitoring the fish faunas of artificial reefs. Fisheries Research 206: 296–308. https://doi.org/10.1016/j.fishres.2018.05.006

Froese R, Pauly D (2019) FishBase. www.fishbase.org

Gratwicke B, Speight MR (2005) The relationship between fish species richness, abundance and habitat complexity in a range of shallow tropical marine habitats. Journal of Fish Biology 66(3): 650–667. https://doi.org/10.1111/j.0022-1112.2005.00629.x

Hagiwara K, Kimura K (2006) Fish fauna of shore waters in Hasama, Tateyama, Boso Peninsula. Memoirs of the National Museum of Nature and Science 43: 351–387.

Hupało K, Majaneva M, Czachur MV, Sire L, Marquina D, Lijtmaer DA, Ivanov V, Leidenberger S, Čiampor F, Čiamporová-Zaťovičová Z, Mendes IS, Desiderato A, Topstad L, Meganck K, Kjærstad G, Lin X, Price B, Stevens M, Ekrem T, Deiner K (2021) An urban Blitz with a twist: rapid biodiversity assessment using aquatic environmental DNA. Environmental DNA 3: 200–213. https://doi.org/10.1002/edn3.152

Isla I (2009) Coastal Zones and Estuaries. In: Isla I, Iribarne O (Eds) Coastal Zone and Esturies. EOLSS Publications, 1–9.

Jerde CL, Wilson EA, Dressler TL (2019) Measuring global fish species richness with eDNA metabarcoding. Molecular Ecology Resources 19(1): 19–22. https://doi.org/10.1111/1755-0998.12929

Johns G, Lee DJ, Leeworthy V, Boyer J, Nuttle W (2014) Developing economic indices to assess the human dimensions of the South Florida coastal marine ecosystem services. Ecological Indicators 44: 69–80. https://doi.org/10.1016/j.ecolind.2014.04.014

Kajiyama M (2009) 30-Year chronicle of stock enhancement effort on chum salmon Oncorhynchus keta based on the release of artificially reared juveniles in Kuriyama River, Chiba, central Japan. Bulletin of the Chiba Prefectural Fisheries Research Center 4: 51–66.

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Molecular Biology and Evolution 30(4): 772–780. https://doi.org/10.1093/molbev/mst010

Kohno H, Kanou K, Yokoo T (2011) A photographic guide to the fishes in Tokyo Bay. Heibonsha, Tokyo, Japan, 374 pp.

Masuda R (2008) Seasonal and interannual variation of subtidal fish assemblages in Wakasa Bay with reference to the warming trend in the Sea of Japan. Environmental Biology of Fishes 82(4): 387–399. https://doi.org/10.1007/s10641-007-9300-z

Meyer R, Ramos M, Lin M, Schweizer T, Gold Z, Ramos D, Shirazi S, Kandlikar G, Kwan W, Curd E, Freise A, Parker J, Sexton J, Wetzer R, Pentcheff N, Wall A, Pipes L, Garcia-Vedrenne A, Mejia M, Moore T, Orland C, Ballare K, Worth A, Beraut E, Aronson E, Nielsen R, Lewin H, Barber P, Wall J, Kraft N, Shapiro B, Wayne R (2021) The CALeDNA program: Citizen scientists and researchers inventory California's biodiversity. California Agriculture 75(1): 20–32. https://doi.org/10.3733/ca.2021a0001

Miya M (2022) Environmental DNA metabarcoding: A novel method for biodiversity monitoring of marine fish communities. Annual Review of Marine Science 14: 161–185. https://doi.org/10.1146/annurev-marine-041421-082251

Miya M, Sado T (2019) Multiple species detection using MiFish primers. In: eDNA Method Standardization Committee (Eds) Environmental DNA Sampling and Experiment Manual Version 2.1. The eDNA Society, Otsu, Japan, 55–92.

Miya M, Higashitarumizu E, Gonoi T, Sunobe T, Mochizuki K (1994a) Fishes of the Boso Peninsula, central Japan-I. Coastal fishes taken by set net off Ainohama, Tateyama. Journal of the Natural History Museum and Institute, Chiba 3: 109–118.

Miya M, Higashitarumizu E, Gonoi T, Sunobe T, Mochizuki K (1994b) Fishes of the Boso Peninsula, central Japan-II. Coastal fishes taken by set net off Chitose, Chikura. Journal of the Natural History Museum and Institute, Chiba 3: 119–128.

Miya M, Toho H, Mochizuki K (1995) Fishes of the Boso Peninsula, central Japan-III. Coastal and deep-sea fishes off Choshi. Journal of the Natural History Museum and Institute, Chiba 3: 195–215.

Miya M, Sato Y, Fukunaga T, Sado T, Poulsen JY, Sato K, Minamoto T, Yamamoto S, Yamanaka H, Araki H, Kondoh M, Iwasaki W (2015) MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: Detection of more than 230 subtropical marine species. Royal Society Open Science 2(7): e150088. https://doi.org/10.1098/rsos.150088

Miya M, Minamoto T, Yamanaka H, Oka S, Sato K, Yamamoto S, Sado T, Doi H (2016) Use of a filter cartridge for filtration of water samples and extraction of environmental DNA. Journal of Visualized Experiments (117): 1–8. https://doi.org/10.3791/54741

Miya M, Gotoh RO, Sado T (2020) MiFish metabarcoding: A high-throughput approach for simultaneous detection of multiple fish species from environmental DNA and other samples. Fisheries Science 86(6): 939–970. https://doi.org/10.1007/s12562-020-01461-x

Nakabo T (2013) Fishes of Japan with pictorial keys to the species, 3rd edn. Tokai University Press, Tokyo, Japan, 2428 pp.

Oka S, Doi H, Miyamoto K, Hanahara N, Sado T, Miya M (2021) Environmental DNA metabarcoding for biodiversity monitoring of a highly diverse tropical fish community in a coral reef lagoon: Estimation of species richness and detection of habitat segregation. Environmental DNA 3(1): 55–69. https://doi.org/10.1002/edn3.132

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2019) Vegan: Community Ecology Package. R package version 2.5–6. https://CRAN.R-project.org/package=vegan

Pawlowski J, Kelly-Quinn M, Altermatt F, Apothéloz-Perret-Gentil L, Beja P, Boggero A, Borja A, Bouchez A, Cordier T, Domaizon I, Feio MJ, Filipe AF, Fornaroli R, Graf W, Herder J, van der Hoorn B, Jones JI, Sagova-Mareckova M, Moritz C, Barquín

J, Piggott JJ, Pinna M, Rimet F, Rinkevich B, Sousa-Santos C, Specchia V, Trobajo R, Vasselon V, Vitecek S, Zimmerman J, Weigand A, Leese F, Kahlert M (2018) The future of biotic indices in the ecogenomic era: Integrating (e)DNA metabarcoding in biological assessment of aquatic ecosystems. The Science of the Total Environment 637: 1295–1310. https://doi.org/10.1016/j.scitotenv.2018.05.002

Saito H (2019) The Kuroshio: its recognition, scientific activities and emerging issues. Kuroshio current: Physical, biogeochemical, and ecosystem dynamics. In: Nagai T, Saito H, Suzuki K, Takahashi M (Eds) Geophysical Monograph Series. American Geophysical Union, 1–10. https://doi.org/10.1002/9781119428428

Senou H, Matsuura K, Shinohara G (2006) Checklist of fishes in the Sagami Sea with zoogeographical comments on shallow water fishes occurring along the coastlines under the influence of the Kuroshio Current. Memoirs of the National Science Museum 41: 389–542.

Sforzi A, Tweddle J, Vogel J, Lois G, Wägele W, Lakeman-Fraser P, Makuch Z, Vohland K (2018) Citizen science and the role of natural history museums. In: Hecker S, Haklay M, Bowser A, Makuch Z, Vogel J, Bonn A (Eds) Citizen Science: Innovation in Open Science, Society and Policy. UCL Press, London, UK, 429–444. https://doi.org/10.2307/j.ctv550cf2.36

Shitamitsu T, Endo S, Mitsui S, Yokochi K, Senou H (2019) List of the fishes of Tateyama Station of the Field Science Center, Tokyo University of Marine Science and Technology, Banda, Tateyama City, Chiba Prefecture, Japan. Bulletin of Kanagawa Prefectural Museum 2019: 121–137. [Nature Science]

Soh W (2003) Transport processes deduced from geochemistry and the void ratio of surface core samples, deep sea Sagami Bay, central Japan. Progress in Oceanography 57(1): 109–124. https://doi.org/10.1016/S0079-6611(03)00054-5

Stecher G, Tamura K, Kumar S (2020) Molecular Evolutionary Genetics Analysis (MEGA) for macOS. Molecular Biology and Evolution 37(4): 1237–1239. https://doi.org/10.1093/molbev/msz312

Steiner F (2014) Frontiers in urban ecological design and planning research. Landscape and Urban Planning 125: 304–311. https://doi.org/10.1016/j.landurbplan.2014.01.023

Taberlet P, Coissac E, Hajibabaei M, Rieseberg LH (2012) Environmental DNA. Molecular Ecology 21(8): 1789–1793. https://doi.org/10.1111/j.1365-294X.2012.05542.x

Taberlet P, Bonin A, Coissac E, Zinger L (2018) Environmental DNA: For Biodiversity Research and Monitoring. Oxford University Press, 253 pp. https://doi.org/10.1093/oso/9780198767220.001.0001

The eDNA Society (2019) Environmental DNA sampling and experiment manual Version 2.1. http://ednasociety.org/eDNA_man-ual_Eng_v2_1_3b.pdf

Thomsen PF, Willerslev E (2015) Environmental DNA– An emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation 183: 4–18. https://doi.org/10.1016/j.biocon.2014.11.019

Tittensor DP, Mora C, Jetz W, Lotze HK, Ricard D, Berghe EV, Worm B (2010) Global patterns and predictors of marine biodiversity across taxa. Nature 466(7310): 1098–1101. https://doi.org/10.1038/nature09329

Yamano H, Sugihara K, Goto K, Kazama T, Yokoyama K, Okuno J (2012) Ranges of obligate coral-dwelling crabs extend northward

as their hosts move north. Coral Reefs 31(3): 663–663. https://doi.org/10.1007/s00338-012-0893-0

Yang S-K, Nagata Y, Taira K, Kawabe M (1993) Southward intrusion of the intermediate Oyashio water along the east coast of the Boso Peninsula I. Coastal salinity-minimum-layer water off the Boso Peninsula. Journal of Oceanography 49(1): 89–114. https://doi.org/10.1007/BF02234011

Yang J, Zhang X, Jin X, Seymour M, Richter C, Logares R, Khim JS, Klymus K (2021) Recent advances in environmental DNA-based biodiversity assessment and conservation. Diversity & Distributions 27(10): 1876–1879. https://doi.org/10.1111/ddi.13415

Supplementary material 1
Table S1

Author: Masaki Miya, Tetsuya Sado, Shin-ichiro Oka, Takehiko Fukuchi

Data type: excel file

Explanation note: eDNA sampling data for the six survey sites.

Link: https://doi.org/10.3897/mbmg.6.80444.suppl1

Supplementary material 2
Table S2

Author: Masaki Miya, Tetsuya Sado, Shin-ichiro Oka, Takehiko Fukuchi

Data type: excel file

Explanation note: List of species detected from MiFish eDNA metabarcoding with raw-read numbers.

Link: https://doi.org/10.3897/mbmg.6.80444.suppl2

Supplementary material 3
Table S3

Author: Masaki Miya, Tetsuya Sado, Shin-ichiro Oka, Takehiko Fukuchi

Data type: excel file

Explanation note:The center of the geographic distribution of each component species calculated from the average latitude of the northern and southern limits from literature records (Nakabo 2013).

Link: https://doi.org/10.3897/mbmg.6.80444.suppl3

Supplementary material 4
Table S4

Author: Masaki Miya, Tetsuya Sado, Shin-ichiro Oka, Takehiko Fukuchi

Data type: excel file

Explanation note: A faunal inventory of the coastal marine fishes of Chiba prefecture compiled from museum collections and literature surveys. Museum acronyms are Natural History Museum and Institute, Chiba (CBM) and its coastal branch (CHMH), National Museum of Nature and Science, Tokyo (NSMT), Kanagawa Prefectural Museum of Natural History (KPM), and Yokosuka City Museum (YCM). All references can be found in the main body of the text.

Link: https://doi.org/10.3897/mbmg.6.80444.suppl4

Supplementary material 5
Supplementary methods

Author: Masaki Miya, Tetsuya Sado, Shin-ichiro Oka, Takehiko Fukuchi

Data type: pdf file

Explanation note: Collection of eDNA samples. eDNA extraction. Paired-end library preparation and sequencing. Data preprocessing and taxonomic assignment. Fish community analysis. Compilation of a faunal inventory.

Link: https://doi.org/10.3897/mbmg.6.80444.suppl5