Summplementary Methods

Ethics statement

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

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. An advanced seminar was conducted on July 4th,2021, to provide an overview on eDNA studies. In addition, a workshop was conducted on organizing safe field trips and avoiding marine accidents, as well as on appropriate water sampling and filtration methods 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 vacation locales.

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 1,000 mL. When the filter was clogged before reaching 1,000 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 diatmeter rubber cap (KOKUGO, Tokyo, Japan), filled with 1 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).

eDNA extraction

The workspace and equipment were thoroughly sterilized prior to DNA extraction. Filtered pipette tips were used and all eDNA-extraction manipulations were conducted in a dedicated room that was physically separated from the pre- and post-PCR rooms 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. An inlet port of each filter cartridge was connected with a 2.0 mL collection tube and the connection between the cartridge and collection tube was tightly sealed with Parafilm. The combined unit was inserted into a 15-mL conical tube, following which the capped conical tube was centrifuged at 6,000 g for 1 min to remove seawater and RNAlater. After centrifugation, the collection tube was discarded, and an aspirator (QIAvac 24 Plus, Qiagen, Hilden, Germany) was used to completely remove the liquid remaining in the cartridge.

The filter cartridge was lysed using proteinase K. Prior to the lysis, PBS (220 μL), proteinase K (20 μL), and buffer AL (200 μL) were mixed and gently pipetted into the cartridge from the inlet port of the filter cartridge. Subsequently, the inlet port was again sealed, and the cartridge was placed in a preheated incubator at 56 °C for 20 min, with constant stirring at a rate of 10 rpm using a rotator (Mini Rotator ACR-100, AS ONE, Tokyo, Japan). After incubation, the film was removed from the inlet port and the port was inserted into a 2 mL tube (DNA LowBind tube, SARSTEDT, Tokyo, Japan) for DNA collection. Thereafter, the combined unit was placed in a 50 mL conical tube and the capped tube was centrifuged at 6,000 g for 1 min to collect the DNA extract.

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.

Paired-end library preparation and sequencing

The workspace and equipment in the pre-PCR area were thoroughly sterilized prior to the library preparation. Filtered pipette tips were used while performing the pre- and post-PCR manipulations in two separate rooms, in order to safeguard against cross-contamination.

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 (5´–ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT NNN NNN GTC GGT AAA ACT CGT GCC AGC–3´), MiFish-U-reverse (5´–GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC TNN NNN NCA TAG TGG GGT ATC TAA TCC CAG TTT G–3´), MiFish-E-forward-v2 (5´–ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT NNN NNN RGT TGG TAA ATC TCG TGC CAG C–3´), MiFish-E-reverse-v2 (5´–GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC TNN NNN NGC ATA GTG GGG TAT CTA ATC CTA GTT TG–3´), MiFish-U2-forward (5´–ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT NNN NNN GCC GGT AAA ACT CGT GCC–3´), and MiFish-U2-reverse (5´–GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC TNN NNN NCA TAG GAG GGT GTC TAA TCC CCG TTT G–3´). 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 six random bases (Ns) present in the middle of these primers enhanced cluster separation on the flow cells during initial base-call calibrations on the MiSeq platform.

The 1st PCR consisted of 35 cycles with a 12 µL reaction volume containing 6.0 µL 2 × KAPA HiFi HotStart ReadyMix (KAPA Biosystems, MA, USA), 2.8 µl of a mixture of the four MiFish primers in a volume ratio of 2:1:1 (U:E:U2 forward and reverse primers; 5 µM), 1.2 µL sterile distilled H2O and a 2.0 µL eDNA template. To minimize PCR dropouts during the 1st PCR (Doi et al. 2019, Miya et al. 2020), eight technical replicates were performed for the same eDNA template using a strip of eight tubes (200 µL). The thermal cycle profile after an initial 3 min denaturation at 95 °C was as follows: denaturation at 98 °C for 20 s, annealing at 65 °C for 15 s, and extension at 72 °C for 15 s, with a final extension at the same temperature for 5 min. A 1st PCR blank (1 B) was also prepared during this process, in addition to EB. However, this experiment was not performed with eight replicates and only a single tube was used for each of the two blanks (EB and 1 B) to minimize the cost.

After completing the 1st PCR, an equal volume of PCR products was pooled from each of the eight replicates in a single 1.5 mL tube. The pooled products were purified using a GeneRead Size Selection kit (Qiagen, Hilden, Germany) following the manufacturer's GeneRead DNA Library Prep I Kit protocol. This protocol repeats the column purification twice to completely remove the adapter dimers and monomers. Subsequently, the purified target products (ca. 300 bp) were quantified using TapeStation 2200 (Agilent Technologies, Tokyo, Japan), 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, 1 B), the 1st PCR products were purified in the same manner, but the purified PCR products were not quantified. Instead, they were diluted according to an average dilution ratio for the positive samples, following which the diluted products were used as templates for the 2nd PCR.

For the 2nd PCR, the following two primers were used to append dual-index sequences (eight nucleotides indicated by Xs) and flow cell binding sites for the MiSeq platform (5′ ends of the sequences before eight Xs): 2nd-PCR-forward (5′–AAT GAT ACG GCG ACC ACC GAG ATC TAC ACX XXX XXX XAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC T–3′); and 2nd-PCR-reverse (5′–CAA GCA GAA GAC GGC ATA CGA GAT XXX XXX XXG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT–3′).

The 2nd PCR was performed over 10 cycles with a 15-µL reaction volume containing 7.5 µL 2 × KAPA HiFi HotStart ReadyMix, 0.9 µL each primer (5 µM), 3.9 µL sterile distilled H2O and 1.9 µL template (0.1 ng/µL with the exceptions of the three blanks). The thermal cycle profile after an initial 3 min denaturation at 95 °C was as follows: denaturation at 98 °C for 20 s, annealing and extension combined at 72 °C (shuttle PCR) for 15 s, with the final extension at the same temperature for 5 min. A 2nd PCR blank (2 B) was also prepared during this process, in addition to EB and 1 B. 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 a 2% E-Gel Size Select agarose gel (Invitrogen, CA, USA). The target amplicons (~ 370 bp) were excised by retrieving them from the recovery wells using a micropipette. The concentration of the size-selected libraries was measured using a Qubit dsDNA HS assay kit and a Qubit fluorometer (Life Technologies, CA, USA), diluted to 10.0 pM with HT1 buffer (Illumina, CA, USA), 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.

Data preprocessing and taxonomic assignment

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) according to the following steps:

1) Forward (R1) and reverse (R2) reads were merged by aligning the two reads using the *fastq merge pairs* command. During this process, 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) were discarded; 2) primer sequences were removed from those merged reads using the *fastx truncate* command; 3) those reads without the primer sequences underwent quality filtering using the *fastq filter* command to remove low quality reads with an expected error rate of >1% and too short reads of <120 bp; 4) the preprocessed reads were dereplicated using the *fastx uniques* command and all singletons, doubletons and tripletons were removed from the subsequent analyses to avoid false positives following the recommendation by the author of the program (Edgar 2010); 5) the dereplicated reads without single- to tripletons were denoised using the unoise3 command to generate amplicon sequence variants (ASVs) that removed all putatively chimeric and erroneous sequences (Callahan et al. 2017); 66) the ASVs were rarefied to the minimum read number (67,618); and 7) finally, ASVs were subjected to taxon assignments to species names (molecular operational taxonomic units; MOTUs) using the *usearch global* command with a sequence identity of >98.5% with the reference sequences (two nucleotide differences allowed) and a query coverage of ≥90%.

PMiFish tentatively assigned “U98.5” labels before the corresponding species names with the highest identities (e.g., U98.5 *Pagrus major*) and they were subjected to clustering at the level of 0.985 using the *cluster smallmem* command. 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, U98.5 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 and U98.5 MOTUs were assembled, all reference sequences were added from that family,

書式を変更: フォント : 斜体

書式を変更: フォント : 斜体

書式を変更: フォント : 斜体

書式を変更: フォント : 斜体

書式を変更: フォント : 斜体

削除: Those

削除: ASVs with the sequence identities of 80–98.5% were tentatively assigned “U98.5” labels before the corresponding species names with the highest identities (e.g., U98.5 Pagrus major)

書式を変更: フォント : 斜体

書式を変更: フォント : 斜体

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. The distances were calculated using pairwise deletion of gaps and among-site rate variations modelled with gamma distributions (shape parameter = 1). Furthermore, bootstrap resampling (n = 100) was performed to estimate the statistical support for the internal branches of the NJ tree and midpoint rooting was performed on the resulting NJ tree.

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...).

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 habitat 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 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, we calculated the center of the geographic distribution for each detected species 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).

Compilation of a faunal inventory

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 (Table S1). Five collections were selected from neighboring natural history museums to represent coastal marine fish fauna in Chiba Prefecture: the 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 pure freshwater fishes were manually excluded (Table S1). For the literature surveys, eight core references were designated (Miya et al. 1994a, 1994b, 1995, Hagiwara and Kimura 2006, Kohno et al. 2011, Nakabo 2013, Aoki et al. 2016, Shitamitsu et al. 2019) and coastal marine fish that appeared to occur in the region were extracted (Table S1).

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: 569–580. https://doi.org/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: 2030–2039. https://doi.org/10.1111/j.1095-8649.2012.03464.x

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

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

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

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.

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Molecular Biology and Evolution 30: 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: 387–399. https://doi.org/10.1007/s10641-007-9300-z

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

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

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: 939–970. https://doi.org/10.1007/s12562-020-01461-x

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 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 Natural History

Museum and Institute, Chiba 3: 119–128.

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, 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: 150088. https://doi.org/10.1098/rsos.150088

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

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. Available from: https://CRAN.R-project.org/package=vegan.

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 (Nature Science) 2019: 121–137. https://doi.org/10.32225/bkpmnh.2019.48_121

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

The_eDNA_Society (2019) Environmental DNA sampling and experiment manual Version 2.1. Available from: http://ednas ociety.org/eDNA_man-ual_Eng_v2_1_3b.pdf.