Supplementary Material to *Dispersal and survival of captive-reared threatened fishes in a Tonle Sap Lake reserve* by T. Campbell et al.

eDNA Laboratory Methodology

Sample Process
1.4.2
Sample barcodes were recorded and assigned a well within the 96 well plate or numbered extraction tube. A customized one-ton arbor press along with a removable leather punch was used to open the plastic casing of each filter. Once plastic casing was cut, sample barcodes were recorded and assigned a well within the 96 well plate or numbered extraction tube. The whole filter was removed and transferred to the extraction plate/tube using sterilized tweezers inside a laminar flow hood. The removable leather punch was sterilized between each eDNA filter. Plates or tubes were immediately processed or stored in -20C until the extraction process could be performed.

Extraction
2.2.3
Genomic DNA from samples was extracted using the DNeasy PowerLyzer PowerSoil Kit (Cat# 12855-100) according to the manufacturer's protocol. Genomic DNA was eluted into 100µl and frozen at -20°C.

PCR
3.18.3
Forward Primer: GTCGGTAAAACTCGTGCCAGC

Reverse Primer: CATAGTGGGGTATCTAATCCCAGTTTG

Primer notes:

Primer reference: Miya et al 2015

Portions of hyper-variable regions of the mitochondrial 12S ribosomal RNA (rRNA) gene were PCR amplified from each genomic DNA sample using the MiFishUF and MiFishUR primers with spacer regions. Both forward and reverse primers also contained a 5' adaptor sequence to allow for subsequent indexing and Illumina sequencing. PCR amplification was performed in replicates of six and all six replicates were not pooled and kept separate. Each 25 µL PCR reaction was mixed according to the Promega PCR Master Mix specifications (Promega catalog # M5133, Madison, WI) which included 12.5ul Master Mix, 0.5 µM of each primer, 1.0 µl of gDNA, and 10.5 µl DNase/RNase-free H2O. DNA was PCR amplified using the following conditions: initial denaturation at 95C for 3 minutes, followed by 45 cycles of 20 seconds at 98C, 30 seconds at 60C, and 30 seconds at 72C, and a final elongation at 72C for 10 minutes. Added 11/2019.

Gel
4.1.1
To determine amplicon size and PCR efficiency, each reaction was visually inspected using a 2% agarose gel with 5µl of each sample as input.

PCR Amplicon Cleanup
5.1.1
Amplicons were then cleaned by incubating amplicons with Exo1/SAP for 30 minutes at 37C following by inactivation at 95C for 5 minutes and stored at -20C.

Barcoding PCR
6.1.1
A second round of PCR was performed to complete the sequencing library construct, appending with the final Illumina sequencing adapters and integrating a sample-specific,12-nucleotide index sequence. The indexing PCR included Promega Master mix, 0.5 µM of each primer and 2 µl of template DNA (cleaned amplicon from the first PCR reaction) and consisted of an initial denaturation of 95 °C for 3 minutes followed by 8 cycles of 95 °C for 30 sec, 55 °C for 30 seconds and 72 °C for 30 seconds.

PCR Normal Pool
8.1.1
Final indexed amplicons from each sample were cleaned and normalized using SequalPrep Normalization Plates (Life Technologies, Carlsbad, CA). 25µl of PCR amplicon is purified and normalize using the Life Technologies SequalPrep Normalization kit (cat#A10510-01) according to the manufacturer's protocol. Samples are then pooled together by adding 5µl of each normalized sample to the pool.

Sequencing
9.7.1
Sample library pools were sent for sequencing on an Illumina NovaSeq 6000 (San Diego, CA) at the Texas A&M Agrilife Genomics and Bioinformatics Sequencing Core facility using the SP Reagent Kit v1.5 (500 cycles) (cat# 20028402). Necessary quality control measures were performed at the sequencing center prior to sequencing.

Bioinformatics
10.11.2
Raw sequence data were demultiplexed using pheniqs v2.1.0 [1], enforcing strict matching of sample barcode indices (i.e, no errors). Cutadapt v3.4 [2] was then used remove gene primers from the forward and reverse reads, discarding any read pairs where one or both primers

(including a 6 bp, fully degenerate prefix) were not found at the expected location (5') with an error rate < 0.15. Read pairs were then merged using vsearch v2.15.2 [3], discarding resulting sequences with a length of < 130 bp, > 210 bp, or with a maximum expected error rate [4] > 0.5 bp. For each sample, reads were then clustered using the unoise3 denoising algorithm [5] as implemented in vsearch, using an alpha value of 5 and discarding unique raw sequences observed less than 8 times. Counts of the resulting exact sequence variants (ESVs) were then compiled and putative chimeras were removed using the uchime3 algorithm, as implemented in vsearch. For each final ESV, a consensus taxonomy was assigned using a custom best-hits algorithm and a reference database consisting of publicly available sequences (GenBank [6]) as well as Jonah Ventures voucher sequences records. Reference database searching used an exhaustive semi-global pairwise alignment with vsearch, and match quality was quantified using a custom, query-centric approach, where the % match ignores terminal gaps in the target sequence, but not the query sequence. The consensus taxonomy was then generated using either all 100% matching reference sequences or all reference sequences within 1% of the top match, accepting the reference taxonomy for any taxonomic level with > 90% agreement across the top hits.

References:

1. Galanti, L.; Shasha, D.; Gunsalus, K.C. Pheniqs 2.0: accurate, high-performance Bayesian decoding and confidence estimation for combinatorial barcode indexing. *BMC Bioinform*. **2021**, *22*, 359.
2. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **2011**, *17*, 10–12.
3. Rognes, T.; Flouri, T.; Nichols, B.; Quince, C.; Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **2016**, *4*, e2584.
4. Edgar, R.C.; Flyvbjerg, H. Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* **2015**, *31*, 3476–3482.
5. Edgar, R.C. UNOISE2: Improved error-correction for Illumina 16S and ITS amplicon sequencing. *BioRxiv* **2016**, 081257. https://doi.org/10.1101/081257.
6. Benson, D.A.; Karsch-Mizrachi, I.; Lipman, D.J.; Ostell, J.; Wheeler, D.L. GenBank. *Nucleic Acids Res.* **2005**, *33* (suppl. 1), D34–D38.