Investigating non-lethal eDNA sampling technique to detect cryptobenthic reef fishes from Vava’u, Kingdom of Tonga

Marta C. Gómez-Buckley1, Ramón Gallego3, (include Karen an Hau here), Luke M. Tornabene1, Ryan P. Kelly2, Raymond M. Buckley1

1 School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA, 2 School of Marine and Environmental Affairs, University of Washington, Seattle, WA, United States of America, 3Northwest Fisheries Science Center, Seattle, WA.

**ABSTRACT**

Cryptobenthic reef fishes (CRF) are an important and often overlooked faunal component of coral reef ecosystems. Their cryptic nature and very small size (average < 20 mm TL), makes them very difficult to assess using visual transect methods. The current method to determine the species composition of CRF is to use enclosures to collect the fishes after an ichthyocide or anesthetic is applied. This study is the first to do comparative collections of underwater eDNA and corresponding samples of the actual fishes collected from the same habitat. The objective of this study was to determine if the diversity of unique DNA sequences from eDNA sampling would reflect the CRF species composition from anesthetic collections. Assessing CRF species composition using eDNA samples would be a faster and non-lethal method, and it would allow participation of citizen scientists in local communities through collection of eDNA water samples. CRF were collected between December of 2018 and January 2019 in Vava’u, Tonga. A total of 596 CRF were collected from shallow (2-5 m) coral heads and coral rubble habitats using quinaldine solution as fish anesthetic, and an underwater airlift sampling device. Before the CRF were collected, modified 200 ml syringes were used to extract water and sequester samples from the coral habitats. A total of 50 eDNA samples were collected at the same stations as CRF. A total of 30 CRF species were observed from the total 596 CRFs collected using visual identification. In contrast, a total of four species of CRFs were detected from the eDNA samples. A low number of reads available for analysis after extraction procedure of the eDNA samples may be one of the main roadblocks in the ability to detect CRFs.

Keywords: [list about six, excluding words in the title]

# INTRODUCTION

The diversity of life that we know coral reef ecosystems harbor is vast. Despite our efforts to diminish this diversity (the Anthropocene), we keep finding increasing numbers of new species in all the known branches of the tree of life whenever we look closer. The smallest invertebrate and vertebrate inhabitants in coral reef ecosystems have been mostly overlooked when performing traditional reef biodiversity surveys. The need to get an accurate census of life in coral reef ecosystems is becoming an urgent issue due to the imperiled status of the shallow reef building corals that house most of these ‘less visible’ reef fauna waiting to be discovered and named.

The smallest fish inhabitants in coral reefs have been termed cryptobenthic reef fishes (CRFs). The most recent definition for this group of coral reef fishes comes from Brandl et al. (2018), which stablished the core group of CRFs “*as those belonging to a family with more than 10% of species being less than 50 mm in body size*” as adults. Their very minute sizes (1-6 cm total length as adults) and erroneously perceived relative small biomass when compared to larger reef fish species (Brandl et al., 2019) may not carry the same level of appreciation as ‘charismatic’ top reef predators, or other reef fishes detectable through traditional surveys, but thanks to recent research efforts on CRFs at the taxonomic and ecological level, we are starting to understand their importance as a critical component in coral reef ecosystems.

CRFs are found in all microhabitats around coral reef ecosystems and often go undetected by divers doing regular surveys of larger reef fishes. Despite their small sizes and their small biomass compared to the larger fishes on the reefs, CRFs contribute tremendous percentages of overall reef biodiversity and energy transfer in the form of available prey items for larger reef inhabitants because of their very high turnover rates (Brandl et al., 2019). CRFs are recognized to represent about 40% of the entire diversity of fishes present in coral reef ecosystems, and about 50% of the abundance by numbers (Depzynski and Bellwood, 2003, 2005, Brandl et. al. 2018). Understanding basic questions of systematics and biodiversity, as well as community assemblages, trophic dynamics, speciation, and habitat utilization, of this group of fishes will lay the foundation for understanding the functioning of the entire reef ecosystem and the evolutionary processes that create and facilitate biodiversity.

Because of CRFs small sizes and cryptic behavior, traditional survey methods are not a feasible for censusing the numbers, or species composition of CRFs. The only method available to answer those questions is by collecting specimens from the reef habitats and euthanizing them. This method requires in situ underwater collections involving ichthyocides, that will slow down the fishes for collections and will eventually euthanize them. To add to this challenge, taxonomic identification by visual morphological analysis requires deep taxonomic expertise, is time consuming, and when not conclusive, additional molecular analysis is required. Furthermore, after the anesthetic agent is spread and CRFs are collected, it is possible that some of the more secretive species that tend to hide deep into the matrix of coral heads or deep into the crevices of coral rubble may be missed.

Environmental DNA is an emerging sampling technique for surveying biodiversity. eDNA can be defined as any molecular material that has shed from living organisms, and that remains in the environment and can be detected when a sample is collected. The nature of the sample can present many origins and be very varied in structure (i.e. from air samples collected to detect microorganisms, spores, or pollen, to scat left behind by animals, and water samples collected from any type of aquatic environments). Because the molecular material that sheds from the living organisms can be free DNA or RNA material or still be incased in a cell membrane or cell wall, the persistence in the environment can vary tremendously and it is also dependent on the surrounding environmental conditions. Furthermore, it may be logical to assume that larger organisms will shed larger amounts of DNA material than smaller organisms, so the material shed from smaller fish like it is the case in this study, could likely be detected easier if collected closer to the bottom environment where they reside as adults.

The overall objective of this study was to determine if eDNA sampling would be a feasible way to survey the species richness of CRFs when working in remote tropical locations with no laboratory facilities and from underwater collections, with specially designed equipment. We specifically compared the CRFs’ species richness using eDNA water samples collected *in situ* from delimited coral reef microhabitats, with the actual CRF specimens that were collected and identified from the same delimited coral reef microhabitats where the eDNA samples had previously been collected. Because this type of study had not been done before, we decided to test which amount of volume would contain detectable amounts of CRFs and took 400 ml and 200 ml eDNA water samples. This was decided based on the fact that

Should talk about the COI as the universal marker for and also the fact that most eDNA studies concentrate on either detecting one or two fishes (endangered or invasive spp), or a general sweep of them, but no studies have tried to concentrate on detecting small fishes that present high diversity.

# METHODS

***Sampling Location***

All samples were collected during a December 2018- January 2019 expedition to Vava’u, the northern most group of islands in the Tongan Archipelago, located 23 degrees south of the Equator. Four sites (Figure 2) were identified during prior expeditions as suitable for the study, and characterized by flat or low-slope profile bottoms, with a relatively homogenous mix of live coral bommies, dead coral rook, coral rubble and sand substrates, at depths rages of 3-7m. All sites are characterized as protected from high swell or high currents under regular weather conditions. Two of the sites, Lotuma Island (18°39'43.0"S 174°00'32.0"W) and Mystic Sands (18°39'53.9"S 174°00'57.7"W) are in close proximity (0.8 km) and near populated areas, while Afo Island (18°42'30.1"S 173°59'47.3"W) and Euakafa Island (18°45'14.7"S 174°02'12.9"W) are separated from the other two sites by about 8 and 11km respectively, and about 6 km apart from each other. The latter two sites/islands are unpopulated, with Euakafa Island being the farthest away from the main populated areas. All sites were located outside of the planned or established Special Management Areas to avoid interference with the Tongan Government efforts to restore and sustain village fishery resources.

***Field sampling and processing procedures of eDNA samples***

Three different microhabitat types were selected for this study based on pilot study and prior sampling trips: Live coral bommies (LC), dead coral rock (DC), and coral rubble (CR). Sand habitats were not included because of the near absence of CRFs when compared to the other microhabitat types. The selected microhabitats were sampled at the above-mentioned stations. To standardize CRF collections from live coral heads, only structurally similar branching *Porites sp.* and *Montipora sp.* coral heads were selected. These two species are present at all the sites as they are resilient corals species that can grow fast and colonize areas of substrate that have experience disturbances after natural or anthropogenic impacts. Constrains of time, boat availability, state the of coral degradation, and distance of some of the sites from the base of operation, dictated the number of replicates for each microhabitat type at each site and depth that were sampled. A minimum of three replicate samples from each microhabitat type and from each site was attempted.

Environmental DNA (eDNA) and CRFs’ samples were collected underwater using standard SCUBA gear. When a suitable microhabitat (i.e., coral head or coral rubble) was selected, eDNA samples were collected at the substrate surface, or within the interstices of the coral rubble or the coral heads, using 200 ml modified syringes. Three technical replicates were be collected of 2 x 200 ml each from each microhabitat. The two syringes for each replicate will be put in three separate containers. Right after the eDNA samples are collected, the CRFs from that same microhabitat were collected using anesthetic stations as described below. The eDNA samples were placed in a cooler when brought to the surface and each technical replicate (400 ml), were be filtered as soon as possible after collection, using a peristaltic pump powered by a hand drill, through a sterile 250 ml disposable filter funnel (Figure 3c) with a 0.45 Micron, 47 mm diameter nitrocellulose mixed ester (MCE) membrane filter (Sterlithech Corporation, Kent, WA). Each replicate sample (filter disc) was preserved in a labeled 2 ml micro-vial with previously filled with 900 μl of Longmire’s solution (ref). In order to fit in the 2ml vials, the filters were folded in half three times with the use of two flat tip tweezers. All the equipment that was reused (syringes, filter funnels, and tweezers) was sterilized between sampling, filtering, and filter storage operations. Sterilization of the equipment was done by soaking it in baths of 10% bleach solution, and thorough rinsing in water, before reusing.

***Field sampling and processing procedures of CRF specimens from anesthetic stations***

All CRF specimens required for Objective 1 will be captured using modified techniques and protocols from (Ahmadia et al. 2018). When a haphazardly selected microhabitat type is located, a cylindrical mesh enclosure (open at the top and the bottom) with lead line at the bottom and float line on the top, is placed over the microhabitat to sequester as best as possible all the CRFs. The enclosure is designed to be able to adapt to the benthic contour of the sampled microhabitat and encompassed an area of approximately 1 m2. A tarp with lead weights in each corner is then placed over the enclosure and anesthetic is delivered inside the enclosure under the tarp. The tarp is gently push down to mix the anesthetic so it would get evenly distribute in the whole volume of the enclosure. For some of the live coral heads, only the covering tarp is used, since some of the coral heads are too massive for the mesh enclosure. After about a minute, when the fish are observed to be responding to the anesthetic, the tarp is lifted slowly and all the CRFs are suctioned into a labeled plastic jar using an airlift device designed specifically for this study (Buckley and Gómez-Buckley, in prep.). Several microhabitats are sampled per dive by two divers working as a team. The number of microhabitats sampled per dive is variable and depends on the time required to locate feasible sites, and on the number of specimens found at each site.

At the end of each dive, the jars with specimens are placed in seawater cooled with jars of ice to keep them from deteriorating in the warm tropical environment. To preserve the specimen’s live coloration requires transporting the CRFs as quickly as possible to the lab station where they are photographed to aid in later detailed morphological analysis. The use of macrophotography lenses makes it possible to document the original coloration of the live specimens as well as all the details that will make it possible to do fin-ray counts, a very important morphological feature for most identification of the CRF specimens. After the fish are photographed, they are humanely euthanized using a heavier dose of the anesthetic solution and then placed in individual 2 ml cryovials, with unique label identifiers, filled with 95% ethanol for later genetic analysis.

All the CRF specimens were identified to the lowest taxonomic level possible by just visual inspection when still fresh and entered into an Excel spreadsheet. Using a photographic box filled with salt water, representative specimens of all the CRFs collected were taken using a Canon XXX camera with a XXX model macro lens attached. Specimens pictures were used post-field operations to confirm identifications in the field, or to determine identification of the unknown ones.

***eDNA lab sample processing***

A dedicated lab with sterilized work surfaces was used to perform DNA extractions from the eDNA samples. Each whole filter disc corresponding to individual sample replicates were extracted and kept as a separate sample for the subsequent amplification and sequencing processing. Prior to the extraction, a ‘bead beating’ processing (Djurhuus et al. 2017) with slight modifications was used to help separate DNA fragments from the filter. The beads were added to the vials that contained the filters in the Longmire’s solution, so the ATL Buffer was used during the bead beating process. The Type of beads, amounts, and give the model of shaker. A DNeasy Blood & Tissue Kit was then used for the extraction following the manufacturer’s protocols.

Extracted raw DNA was sent to Jonah Ventures Lab, Boulder CO, for sequencing. The normal length of specific fish markers is about 600 bp, but we are dealing with degrade DNA in the environment, so Leray primers (Leray et al. 2013) that target a shorter fragment of about 300 base-pair-long from the COI gene will be used. Other markers like 16s and 12s will not be used, because they are not as common in GenBank, and in many cases, they do not give information on higher taxonomic levels than the ones needed for my research question.

***Bioinformatics analysis***

Sequences were processed with the JAMP pipeline (https://github.com/VascoElbrecht/JAMP). Demultiplexing of raw data was done with "iu-demultiplex" v2.3 (https://github.com/merenlab/illumina-utils). Forward and reverse reads from each sample where paired end merged with Usearch v11. Primers were trimmed from both sides and reads of 298 - 328 bp length retained using Cutadapt v1.15. Quality filtering was carried out using expected error filtering (max\_ee=1) as implemented in Usearch v11. For OTU clustering reads where pooled, dereplicated and singletons removed. Reads where clustered using usearch at 3% similarity, and each sample (including singletons) remapped against the generated OTUs using usearch\_global with 97% similarity and maxaccepts=8, maxrejects=256 to generate a OTU table. Taxonomy was assigned using using the latest BOLD database (All Barcode Records on BOLD, BOLD\_web\_hack.R).

More text, etc. In some cases, if the statistical analyses or complex or unusual they may be described in the Methods, but if only very conventional analyses were done (e.g., ANOVA, t-tests, linear regression) they can be indicated in the Results with no need for additional mention in the Methods.

# RESULTS

*CRFs from anesthetic stations*

A total of 683 CRFs were collected from the anesthetic stations from where eDNA samples were taken. The number of CRF species identified from the composite of all the stations was 34. From this, XX were Gobiidae, with other families present in very small numbers (Fig. X – chart of fish silhouettes with proportional areas for each of the families. If possible, insert within the Gobiidae, different silhouettes for the different genera? Or just proportional area boxes).

Within the Gobiidae, the *Eviota* genus made out about half of all CRF species, with ~40% of all CRFs belonging to one single species, *E. sigillata*.

The different microhabitats, live coral heads, dead coral, and coral rubble, were approximately sampled at similar numbers. (table ?)

*eDNA lab processing*

*CRF sequences detected from in situ eDNA samples*

Include figure of rarefaction curves (not enough CRF detected per sampl e)

*eDNA*

eDNA samples detected a wide range of the metazoan community present in the environment. As it was expected, the comparative amount of OTUs assigned to CRFs was minimal.

If there was a lab component and field component listed as sub-sections in the Methods, then the same order and structure should be repeated in the Results.

It is very important to state all the important results in words in this section, rather than relying on the figures and tables to tell the story. Tell the story here, and then use the tables to archive specific results, and the figures to highlight the most interesting and important findings because humans, being such visual animals, instinctively look at the pictures. Avoid sentences such as, “The growth rate and sex ratio data are shown in Figure 1.” This tells the reader nothing, and forces him or her to go look at the Figure and decide what it means. It is much better to make a declarative statement and then refer to the figure parenthetically. For example, “Male copper rockfish were consistently larger than females for a given age, but the sex ratio was increasingly female-biased after age 10 (Fig. 1).

# DISCUSSION

* Issues to be solved (sources of error):
* In lab contamination *Eviota Lachdeberie*. Samples that did not present contamination had nearly zero sequences of other fishes present, which indicates that the presence of contamination for later process samples did not affect/overshadowed/obscured/ the detection of other fish species. no other fishes.
* Outsourcing the PCR and sequencing to a lab, advantages and disadvantages.
* Reed depth during sequencing
* Refine extraction method
* Create a mockup sample with DNA extracted from the fish that were collected from specific anesthetic station and try to match up with eDNA sample collected from that same station. Initial cost of extracting and sequencing, but would resolve the issues with insufficient or erroneous matches in public DNA databanks.
* 200 vs 400 question and bring in Ryan and Ramon’s tide paper and the one on maximizing fish eDNA. and the one on cumulative curves
* Universal primers used.
* Nguyen et al. 2020 mentions on discussion how the “*The co-amplification of numerous non-metazoan taxa lowered the sampling depth of our target organisms (metazoan) and likely prevented the consistent detection of rare taxa*”
* Still detected fish, but very low when compared to what is found on anesthetic stations.
* An important roadblock for biodiversity projects relying on eDNA metabarcoding techniques is the known gaps and errors known to exist on the available databases, used to match the sampled environmental sequences, especially for species in regions that have been less studied. Despite all these pitfalls, eDNA sampling is becoming a widely used sampling technique with more and more adept users hailing it as the future of sampling for biodiversity projects.
* eDNA is a possible way to collect information about CRFs without the destructive implications.
* DiBattista et al. (2017) collected eDNA water samples from the surface above reef habitats in the Red Sea. They found a significant missing portion of fish species that were known to be in the region. They collected samples of surface water above the reef, and used 16s marker, with resolution only to genus level. Nguyen et al. (2019) conducted visual fish surveys, and compared the results with the fish detected in eDNA samples collected 10-20 cm above the different habitats surveyed. The reads assigned to bony fishes (COI was the maker used) was almost negligible, amounting to a median of 0.07 %. No eDNA studies known to date have collected *in situ* underwater eDNA samples from the matrix of the reef. This approach could make it easier to detect reef fishes in eDNA samples, especially CRFs which are in physical contact with the substrate as one of their behavioral traits. The question is, could eDNA samples be used to inform us about CRFs’ community assemblages? If so, how would it compare with the information we can gather from the use of anesthetic stations? More research is necessary to refine and bring eDNA sampling to the forefront of ecological studies as it holds great potential to be used as a complementary method for biomonitoring, conservation, and biodiversity efforts, as well as for citizen science projects (Deiner et al., 2017).

Starting with the interpretation of your data in light of your hypotheses.

Then compare your data to other studies on the same or related organisms or topics. Then a series of progressively broader paragraphs, ending where the Introduction began.

It is a cheat to introduce your subject with broad and lofty themes of ecology and the meaning of life, but then only narrowly interpret your data and never return the broad themes, thus failing to fulfill the implied promises of the Introduction.

**ACKNOWLEDGMENTS**

Anyone who helped with ideas, data, Excel/R assistance, editorial help, and other professional services. It is important to acknowledge funding sources, and in some cases very specific wording is requested.

**REFERENCES**

**Note the way that they are set up, with hanging indents in the paragraph formatting menu, rather than tabs or hard returns. The format must be fine-tuned for the style manual of the journal, and there is a nearly infinite variety of formatting details. Look it up – do not make it up. Note especially that references to websites (which should be only when absolutely needed) should indicate the entity responsible for the website, the full web-link, and the date accessed. The following is a set of examples; please follow them and check the style manual of Transactions of the American Fisheries Society for any questions.**

Allen, G.H. 1956. Migration, distribution, and movement of Puget Sound silver salmon. Doctoral dissertation. University of Washington, Seattle.

Allen, G.H. 1959. Growth of marked silver salmon (*Oncorhynchus kisutch*) of the 1950 brood in Puget Sound. Transactions of the American Fisheries Society 88:310–318.

Andrews, K.S., P.S. Levin, S.L. Katz, D. Farrer, V.F. Gallucci, and G. Bargmann. 2007. Acoustic monitoring of sixgill shark movements in Puget Sound: evidence for localized movement. Canadian Journal of Zoology 85:1136–1143.

Batschelet, E. 1981. Circular statistics in biology. Academic Press, London.

Beetz, J.L. 2009. Marine survival of coho salmon (*Oncorhynchus kisutch*) in Washington State: characteristic patterns and their relationship to environmental and biological factors. Master’s thesis. University of Washington, Seattle.

Dean, C.A., T.W. Kassler, and G. Mendel. 2010. Genetic Identification of Brook Trout or Brook / Bull Trout Hybrids in Pataha Creek, WA. Washington Department of Fish and Wildlife Report. Available: http://wdfw.wa.gov/publications/01139/wdfw01139.pdf. (November 2012).

Fresh, K.L., R.D. Cardwell, and R.R. Koons. 1981. Food habits of Pacific salmon, baitfish, and their potential competitors and predators in the marine waters of Washington, August 1978 to September 1979. State of Washington Department of Fisheries, Progress Report No. 145.

Healey, M.C. 1991. Life history of Chinook salmon (*Oncorhynchus tshawytscha*). Pages 311-394 *in* C. Groot and L. Margolis, editors. Pacific salmon life histories. University of British Columba Press, Vancouver.

**Table 1.** Caption here.

Table here, below the caption.

Avoid putting boxes around the table – just a line under the headings.

**Table 2.** Caption here.

Table here, below the caption.

**Figure captions.**

List all the figure captions here, with an extra space between them.

Fig. 1**.** Locations of 21 receivers used for within-basin analysis; the size of the circle represents the score of that location on the first principal coordinate. Receivers are numbered in order of decreasing site use (1 = most frequently used, see Table 2), with receivers categorized as deep/offshore in white and shallow/onshore in gray.

Fig. 2. Caption here

**Fig. 1.**

Figures should always be in black and white unless the journal if produced on-line only. Such on-line journals are usually happy to have color but color in print journals is really expensive and they will charge you for it.

Avoid putting boxes around figures, and avoid titles or headings within the figure that are redundant with the captions. Most journals also want you to describe in the caption what the different symbols (open circles, filled diamonds, etc.) or bar colors (white, black, cross-hatched) refer to. I think this is stupid but I’ve lost the battle on many occastions. So, rather than having information in the figure itself (e.g., black bars = males, white bars = females) you’d say so in the caption.

Brainstorm ideas and notes for consideration