**Proposed Title of Paper**

Authors: Megan, Eily, Ryan, Kim

Target Journal: eDNA or similar

**Abstract**

There are many choices that simply must be made when sampling for eDNA analysis, whether a targeted, single-species approach or a community-based, metabarcoding approach. Of the entire workflow from sampling water to bioinformatic analyses, it seems like the choices in molecular methods and downstream bioinformatics are the most important. However, the first steps in the process of collecting water, filtering it, and preserving it are also important and are often overlooked as major decision points. Many researchers use established methods with limited or no testing for their particular research question. Here, we investigate (1) the volume of water filtered and the filter pore size and (2) the preservation method and extraction method of samples with a specific lens on how inhibition and rarity of target interact with these choices to impact the detection of a single targeted assay (via qPCR and ddPCR) and community compositions (via metabarcoding). We find that XXX filtering larger volumes and maximizing total DNA yield during extraction is not always beneficial to detection likely due to the concentration of inhibitors and co-extraction of off-target template. We recommend collecting pilot samples to determine the minimum volume of water required given the rarity of the target and using an extraction method tailored to the taxa of interest. Finally, we note that inhibition is often masked in metabarcoding read proportions and recommend quantifying inhibition in environmental samples before metabarcoding to avoid missing rare taxa or misinterpreting proportions of target taxa.

**Scope**

* Time
  + Wrangling existing information: 1 week
  + ddPCR / qPCR follow up: 1 week
  + PCR + 1 Miseq run + processing: 1 week
  + Analysis + writing + out the door: 2 months
* Costs
  + ddPCR/qPCR: ~$1,000
  + PCR/MiSeq: ~$2,000

**Methods**

* *Samples in hand (see* [*DURIP*](https://docs.google.com/document/d/1mwMBkgRRDI2rjMt3b9j9C5Y6mehgQ81p/edit) *proposal Tables 2/3 for overview)*
  + Homogenized, high template, inside dolphin pen
    - **N = 12**
    - 2 pore sizes (1 um, 5 um), 2 volumes (1 L, 3L), triplicate
    - All preserved in Longmires, extracted via PC
    - Already been quantified via qPCR (\*were they inhibited?)
    - **No difference between pore size, 3x volume is more than 3x target**
    - **Inhibition…?**
  + Non-homogenized, high template, inside dolphin pen
    - **N = 36 (12 sets of triplicates)**
    - All 5 um, 3 L filtered
    - 5 preservatives (self, -80, Longmires, Shield, Later), 3 methods (PC, Qiagen, Zymo), triplicate \*\*but not full matrix so reduced
    - Already been quantified via qPCR (\*were they inhibited?)
    - **-80 was best preservative but is impractical, Shield was best otherwise**
    - **Qiagen seems like best extraction method, PC noticeably lower for target (despite higher total DNA)**
    - **Inhibition…?**
* *To do* 
  + Wrangle existing data
    - Total DNA of all extracts
    - Inhibition level
    - Target DNA (T. turiops) by qPCR
  + ddPCR extracts
    - Run all samples at 1:1, dilution which no longer inhibited, and then one dilution much lower than lowest of all – *goal is to see if ddPCR gives same quantifications in samples that we know to be inhibited and if the math works out for dilutions (before and after inhibition is diluted out)*
    - Potentially multiplex with IPC + Clupea or something else
      * *IPC to double check inhibition and if you can quantify on ddPCR*
      * *Second species-specific assay to get a second quantification for comparing to metabarcoding proportions (see below)*
  + Amplify and sequence
    - MarVer1 so everything should be perfect match (*for now, don’t do COI or more general marker because we should be able to tell based on total DNA of extract vs. target DNA especially if we did 2 single species ddPCR assays how much is junk in the haystack)*
    - Select subset so it is not all of them – proposed:
      * **N = 93 (31 in triplicate)**
      * 12 samples from volume / pore size
        + *Goal here is to* *see if anything interesting comes up in different taxa from pore size or if anything is masked in 3L vs. 1L samples because of the bigger haystack*
      * 10 samples from preservation/extraction
        + Drop RNALater but otherwise all samples

PC is only with all four preservatives

\*definitely need Longmires/PC – to tie to volume/pore size experiment

All 3 extractions for self-preserving, Shield, and -80

* + - * + *Goal here is to say which combination gives you the best recovery of rare species and try to assess if preservation or extraction method introduces species-specific biases?*
      * 9 samples for inhibition testing
        + 9 -- 3 samples known to be inhibited run at 3 dilutions (1:1, dilution where inhibition gone, and way below that)
        + *Goal here is to measure impact of inhibition on proportions after metabarcoding*

**BIG PICTURE CONTRIBUTIONS**

\*Trade-off in filtering larger volumes and inhibition. More needles in a bigger haystack.

\*\*Does ddPCR do a better job with inhibition than qPCR?

\*\*Does the math math when you do dilutions for ddPCR/qPCR?

\*\*How does inhibition affect proportions of species after metabarcoding?

\*How much does preservation and extraction impact results?

\*\*Species specific quantifications

\*\*Metabarcoding proportions

Rare 🡪 Common x Not inhibited 🡪 Inhibited

* fill in matrix with best option for preservation and extraction
* community composition?