Time to get real with qPCR controls: The frequency of sample contamination and the informative power of negative controls in environmental (e)DNA studies

Patrick Hutchins[1], Adam J. Sepulveda[2], and Leah Simantel[3]

[1]US Geological Survey
[2]US Geol Survey
[3]Affiliation not available

July 6, 2021

Abstract

Environmental (e)DNA methods have enabled rapid, sensitive, and specific inferences of taxa presence throughout diverse fields of ecological study. However, use of eDNA results for decision-making has been impeded by uncertainties associated with false positive tests putatively caused by contamination. Sporadic contamination is a process that is inconsistent across samples and systemic contamination occurs consistently over a group of samples. Here, we used empirical data and lab experiments to (1) estimate the sporadic contamination rate for each stage of a common, targeted eDNA workflow employing best practice quality control measures under simulated conditions of rare and common target DNA presence, (2) determine the rate at which negative controls (i.e., "blanks") detect varying concentrations of systemic contamination, (3) estimate the effort that would be required to consistently detect sporadic and systemic contamination. Sporadic contamination rates were very low across all eDNA workflow steps, and, therefore, an intractably high number of negative controls (>100) would be required to determine occurrence of sporadic contamination with any certainty. Contrarily, detection of intentionally introduced systemic contamination was more consistent; therefore, very few negative controls (<5) would be needed to consistently alert to systemic contamination. These results have considerable implications to eDNA study design when resources for sample analyses are constrained.

Target journal : Molecular Ecology Resources, molecular and statistical advances that demonstrate existence of important problems with current procedures.

Title: Time to get real with qPCR controls: The frequency of sample contamination and the informative power of negative controls in environmental (e)DNA studies

Running title: Environmental DNA contamination detection

Authors: Patrick R. Hutchins[1*], Leah N. Simantel[1], Adam J. Sepulveda[1]

Affiliations:

[1]U.S. Geological Survey Northern Rocky Mountain Science Center, Bozeman MT, USA, 59715

*corresponding author: phutchins@usgs.gov

Disclaimer: This draft manuscript is distributed solely for purposes of scientific peer review. Its content is deliberative and predecisional, so it must not be disclosed or released by reviewers. Because the manuscript has not yet been approved for publication by the U.S. Geological Survey (USGS), it does not represent any official USGS finding or policy.

Disclosure Statement: The authors have nothing to disclose

Abstract

Environmental (e)DNA methods have enabled rapid, sensitive, and specific inferences of taxa presence throughout diverse fields of ecological study. However, use of eDNA results for decision-making has been impeded by uncertainties associated with false positive tests putatively caused by contamination. Sporadic contamination is a process that is inconsistent across samples and systemic contamination occurs consistently over a group of samples. Here, we used empirical data and lab experiments to (1) estimate the sporadic contamination rate for each stage of a common, targeted eDNA workflow employing best practice quality control measures under simulated conditions of rare and common target DNA presence, (2) determine the rate at which negative controls (i.e., "blanks") detect varying concentrations of systemic contamination, (3) estimate the effort that would be required to consistently detect sporadic and systemic contamination. Sporadic contamination rates were very low across all eDNA workflow steps, and, therefore, an intractably high number of negative controls (>100) would be required to determine occurrence of sporadic contamination with any certainty. Contrarily, detection of intentionally introduced systemic contamination was more consistent; therefore, very few negative controls (<5) would be needed to consistently alert to systemic contamination. These results have considerable implications to eDNA study design when resources for sample analyses are constrained.

Keywords: environmental DNA; eDNA, contamination, negative control

1| Introduction

Environmental (e)DNA sampling is a rapidly expanding and evolving approach to detect organismal DNA from environmental matrices (e.g., water, air, soil, surfaces; Sepulveda, Hutchins, Forstchen, McKeefry, and Swigris 2020). Technologies in this field have enabled rapid, sensitive, and specific inferences of taxa presence in a variety of contexts and for a variety of purposes (Barnes & Turner, 2016; Bass, Stentiford, Littlewood, & Hartikainen, 2015; Cristescu & Hebert, 2018). However, use of eDNA results for decision-making has been impeded by the uncertainty of eDNA detections since multiple sources of error can give rise to observation and site-level false positives (Darling, Jerde, & Sepulveda, 2021). Decision makers require greater confidence that eDNA detections are not erroneous because there can be high economic, social and political costs associated with these decisions (Sepulveda, Nelson, Jerde, & Luikart, 2020).

Environmental DNA sampling enables detection of short fragments of DNA at extremely low concentrations, but this strength comes at the cost of heightened susceptibility to small amounts of contamination (Sepulveda, Hutchins, et al., 2020). We define eDNA contamination as detections of specific DNA targets that are not attributable to the presence of those DNA targets in the sample matrix prior to sample collection. This definition excludes other instances of false positives in eDNA sampling caused by experimental or assay design flaws or PCR amplification error, such as detection of non-specific or off-target DNA. Contamination is a demonstrated problem; potential for contamination was reported in 6% of targeted eDNA studies from 2008 – 2019, and contamination was attributed to most stages of the eDNA workflow, from the preparation of field sampling supplies to PCR analyses (Sepulveda, Hutchins, et al., 2020).

Those using eDNA sampling, and, more generally, PCR approaches (Borst, Box, & Fluit, 2004; Weyrich et al., 2019) are aware of contamination risk and have developed best practices to prevent contamination in the field (e.g., single-use supplies, bleach sterilization) and in the lab (e.g., separation of low-template vs. high-template DNA work spaces), as described in Goldberg et al. (2016). However, there is less consensus on best practices for detecting eDNA contamination outside of the need to include negative controls, or "blanks" (i.e., sample material lacking target DNA), during field collection and PCR analyses (Bustin et al., 2009; Goldberg et al., 2016). There is still disparity, for instance, in the number of negative controls to include and at which steps of the workflow they should be included. For example, 49% of targeted eDNA studies reviewed in (Sepulveda, Hutchins, et al., 2020) limited negative controls to just laboratory procedures. Furthermore, these measures are often employed with no empirical basis, and their associated uncertainties remain unknown. The lack of information regarding the number of samples required

for detection of contamination events is surprising because it is generally recognized that detection of rare, target DNA in the field requires considerable eDNA sampling effort (Erickson, Merkes, & Mize, 2019).

Here, we used empirical data and lab experiments to explore the occurrence and frequency of sporadic and systemic contamination in a targeted eDNA workflow. We define "sporadic contamination" as a contaminating process that is inconsistently/irregularly applied across samples or replicates (e.g., sample-to-sample crossover) and "systemic contamination" as a contaminating process that is applied consistently/regularly over a group of samples or replicates (e.g., using contaminated reagents). These terms were chosen for clarity in the context of this study, but correspond with the definitions of "sample contaminant" and "general contaminant" given in Borst et al. (2004). We focus on targeted eDNA approaches because they have been more commonly used in applied surveillance programs than metabarcoding approaches. However, the initial metabarcoding eDNA workflow (sample collection through PCR) parallels targeted approaches; thus, our results should be generally applicable. Our objectives were to (1) estimate the sporadic contamination rate for each stage of a targeted eDNA workflow under simulated conditions of rare and common target presence, (2) determine the rate at which commonly employed controls detect varying concentrations of sporadic and systemic contamination, and (3) estimate the effort that would be required to detect sporadic and systemic contamination with varying levels of confidence at each stage of eDNA workflow.

2| Materials and Methods

Our methods are partitioned into investigations of (1) sporadic contamination (Fig. 1a) and (2) systemic contamination (Fig. 1b) at principle steps of a common, targeted eDNA workflow similar to Carim et al. (2015), Laramie, Pilliod, Goldberg, and Strickler (2015), and Minamoto et al. (2021). Our workflow is separated into 4 general steps (see Figure 1): field sample collection, sample concentration (filtration in this study), DNA extraction, and DNA amplification (PCR). As these nested steps cannot be isolated from one another, we cannot unambiguously discern contamination that occurred in an earlier process from one that occurred at a later process.

We analyzed all samples for a DNA sequence of the elongation factor-1 alpha (EF1a) gene in all species of the Salmonidae family (assay design and validation are described in Appendix 1). We chose this DNA target because the facility where laboratory procedures were performed frequently handles samples from salmonids and salmonid waters, thus there is potentially contaminating material at that facility. The log-linear slope, intercept, R^2, and efficiency of the assay across a 12-times replicated seven-point standard curve (4 to $4e^6$ gene copies) were -3.30, 37.95, 0.99, and 1.01, respectively. The limit of detection and limit of quantification ($\pm$ standard error), as determined following the methods of Forootan et al. (2017), were 2.08 (1.05) and 9.41 (1.09), respectively.

2.1| Laboratory Conditions

All laboratory processes were performed at the US Geological Survey's Northern Rocky Mountain Science Center in Bozeman, Montana (USA). Quality control best-practices in place for eDNA workflows included: unidirectional workflow from low to high DNA concentration processes, spatial separation of processes, pre- and post-process surface decontamination with 10% bleach, 70% ethanol, and UV, use of sterile work hoods for reagent preparation and DNA extraction, and single-use materials for sample collection. Laboratory protocols were carried out by a single individual.

Water samples were filtered in the laboratory through 1.5-µm glass microfiber filters (Whatman catalog #1827047) attached to sterilized, filter-funnel cups and a vacuum filter apparatus using a peristaltic pump (Geotech Environmental Equipment, Denver, Colorado, USA). Six samples were processed at a time and filter funnels were spaced 10 cm apart. Each filter was folded in half three times and placed into a buffered solution of protease K and digested for at least eight hours at 56 °C. All DNA extraction methods were carried out in a sterilized laminar flow hood using Qiagen DNeasy Blood and Tissue Kits (Qiagen.com, catalog #69506) according to the manufacturer's instructions except that samples were digested in Qiagen Investigator Lyse & Spin Baskets (Qiagen.com, catalog #19597).

Amplification was carried out on a CFX96 Touch (Bio-Rad, Hercules, California) with 20 µL volumes on 96-well Bio-Rad PCR plates (#HSP9601) sealed by hand with Bio-Rad Microseal 'B' optical sealing film (#MSB1001). Reactions included 10-µL Qiagen Quantitect Probe PCR Master Mix (#204343), 0.5 µM of each of the forward and reverse primers, 0.25 µM of the probe, 4 µL of sample DNA extract, and sterile water to achieve 20 µL. The thermal cycle used was 15 minutes at 95 °C followed by 40 cycles of 15 seconds at 95 °C and 1 minute at 60 °C. Eight qPCR replicates were performed for each sample. Raw fluorescence data was baseline corrected according to Patrone, Romsos, Cleveland, Vallone, and Kearsley (2020) and a threshold fluorescence value above the noise floor of early cycles was computed for all samples using custom functions in R v4.0 (R-Core-Team, 2014). A positive detection was defined as any sample that crossed this common fluorescence threshold before cycle 40.

2.2| Sporadic contamination

Field sample collection . Weekly water field samples were collected from 7 locations with high densities of salmonid fishes on the Madison and Yellowstone rivers, Montana, USA from June through September 2020 as part of a salmonid parasite monitoring program (Hutchins et al., 2021). Field samples were collected using sterile Whirl-Pak bags (Nasco, Fort Atkinson, Wisconsin, USA). Onsite, but prior to field sample collection, we poured 250 mL of reverse osmosis (RO) water into a sterile, single-use Whirl-Pak bag, exposed the bag to the field environment for 10 seconds, and then closed the bag. These field negative control samples (n=120) were transported alongside field samples to the laboratory for further processing that occurred in parallel with field samples (see Figure 1a).

Filtration . We performed two treatments to assess contamination risk when negative samples are co-filtered with positive samples in the laboratory. First, we filtered 250 mL of RO water sample (n=120) in succession. Second, each 250-mL RO water sample (n=60) was followed by a 250-mL sample from water containing rainbow trout DNA (n=60) made by thawing fish carcasses (˜250 g wet mass) in 20 L of room temperature water for 12 hours. Fish carcasses were removed, and the water was mixed via magnetic stir bar while dispensing filtration aliquots.

DNA extraction. We performed two treatments to assess contamination risk when negative samples are co-processed with positive samples. First, non-spiked extractions were performed using only kit reagents (n=120). Second, non-spiked extractions (n=60) were alternated with extractions (n=60) spiked with $1e^6$ synthetic gene copies of template DNA.

PCR . We performed three treatments to assess differences in contamination risk when no-template controls (NTCs) are co-processed with replicates from positive sample material. The first treatment was 960 NTC qPCR reactions run without any co-processed samples that contained our target DNA ("all negative"). The second treatment consisted of 480 NTC reactions run alongside 480 positive reactions spiked with $4e^3$ synthetic gene copies ("mixed"). Spiked and NTC reactions were arranged in a checkerboard pattern on ten 96-well plates. The third treatment included 900 NTC reactions run on plates that contained a single six-point standard curve series of our target DNA in column 12 of each 96-well PCR plate ("standard curve"). Differences in reaction sample size among the three treatments reflect the constraints of a 96-well plate.

2.3| Systemic Contamination

We evaluated systemic contamination in negative controls introduced during filtration, extraction, or PCR workflow steps. We intentionally contaminated either the sample water itself or a reagent used during the procedure for that workflow step with either a "high" or a "low" contamination treatment (see Figure 1b).

Filtration. We pipetted 1.6 and 0.16 mL of water incubated with rainbow trout (described previously) into 250 mL of RO water for the high (n = 60) and low (n = 60) contamination treatments, which equated to approximately $1e^1$ and $1e^0$ gene copies, respectively, in each qPCR replicate.

DNA Extraction. We carried out the extraction as described above except that we did not add any sample material, and we intentionally contaminated the 90% ethanol that was used in the protocol with $6e^3$ and

$6e^2$ synthetic gene copies for the high (n = 60) and low (n = 60) contamination treatments. This equated to approximately $1e^2$ and $1e^1$ gene copies, respectively, in each qPCR replicate.

PCR. We used sterile water as the sample material for NTCs except that we intentionally contaminated the assay master mix with synthetic gene copies. The average final amount of contaminating gene copies per each qPCR reaction for the high (n = 480) and low (n = 480) treatments were $1e^0$ and $1e^{-2}$ gene copies, respectively.

2.4| Statistical Analysis

We used Bayesian multi-scale occupancy models to account for false negatives and to evaluate workflow and treatment effects on the probabilities of detecting contaminating DNA in PCR replicates (p), samples (ϑ), or sample sets (Ψ) in the sporadic and systemic contamination datasets (R version 4.0., msocc package, Stratton, Sepulveda, and Hoegh 2020). For the purposes of modeling these probabilities in the amplification workflow step, a "sample" was defined as 8 PCR reactions (i.e. PCR replicates) within a distinct column on 96-well plates (Figure 1). For the sporadic contamination dataset, we compared support of a null model to models that included the workflow step, the parallel processing components, or their interaction as covariates of Ψ , ϑ and/or p . For the systemic contamination dataset, we compared support of a null model to models that included the workflow step, the treatment level, or their interaction as covariates of Ψ , ϑ and p . We used the widely applicable information criteria (WAIC; Watanabe 2010) to compare support for models fitted with and without covariates; models with lower WAIC values are favored (Gelman, Hwang, & Vehtari, 2014). We then computed estimates of the derived parameters Ψ, ϑ , and$p$ for the most favored model. These estimates and their standard errors were computed using a single Markov chain containing 10,000 iterations (excluding the first 1000 warm-up iterations, which were discarded as burn-in). Convergence was assessed using traceplots provided in the msocc package R Shiny web application.

Next, we evaluated how sample size (i.e., the number of water samples and the number of PCR replicates) influenced the precision of estimates (msocc package, msocc_sim(); Stratton et al. 2020). We used the estimates of ψ, ϑ, and p from the most supported models to simulate detection data; we varied the number of samples collected at each sampling event and the PCR replicates analyzed per sample. We then replicated this process 100 times and assessed the sample sizes at which the average width of the credibility intervals stabilized. This provided insight about the point of diminishing returns, beyond which increasing sample size provides little benefit.

We then used the derived estimates of ϑ and p from the most favored sporadic and systemic contamination models to estimate ${\vartheta_{50}}^*$,${\vartheta_{95}}^*$, ${p_{50}}^*$, and${p_{95}}^*$, the number of water samples or PCR replicates required to have 50% or 95% probabilities of detecting target DNA, conditional that it is present (Sepulveda, Amberg, & Hanson, 2019). Our primary interests were estimating ${\vartheta}^*$ and ${p}^*$ for the sample collection, filtration and extraction workflow steps andp^* for the PCR workflow step. We used the following equations for these estimates, where n is the number of samples or replicates, ϑ and p are the median, lower 95% credible interval (CI) or upper 95% CI widths of the conditional probabilities of occurrence at the sample and replicate levels:

$$\theta^* = 1 - (1-\theta)^n;$$

$$p^* = 1 - (1-p)^n.$$

3| **Results**

3.1| Sporadic contamination

Naïve results . Amplification of contaminating DNA was very rare. Thirteen of 810 PCR reactions amplified, and none of these thirteen samples had more than one PCR replicate amplify (Table 1). There was only one PCR replicate with a quantification cycle (Cq) <35, which crossed the threshold at cycle 20.87. The copy number estimate for this outlier was $1.60e^5$ gene copies, and the mean (± standard deviation) Cq and copy number estimate of the other 12 replicates were 38.43 (1.37) and 1.29 (2.13), respectively. Amplification was more frequent for 'all negative' parallel processing steps than for mixed or standard curve parallel processing

steps (Table 1). However, this pattern was confounded by sample size as the number of samples associated with 'all negative' parallel processing was much more than other steps.

Modeled results . The most supported model was Ψ (workflow step*parallel processing), ϑ(workflow step*parallel processing),p (.), with a WAIC value that was 396 WAIC units lower than any other model. Mean posterior estimates of p were near zero because few PCR replicates amplified. Given the low p estimates, CI widths of Ψ and ϑ estimates were large (0.47 – 1.00) and posterior mean estimates of Ψ and ϑ estimates were uninformative (Fig 2a). Post-hoc power analyses indicate that the CI width for Ψ , ϑ, and p would not decrease with more PCR replicates. The points of diminishing returns for Ψ , ϑ, and $p$confidence interval widths were 5, 1, and 8 PCR replicates per sample.

Based on the derived mean estimate of p , 157 PCR replicates are needed for $p_{0.50}$* and 678 PCR replicates are needed for $p_{0.95}$* (Fig. 3). These results underscore that detection of sporadic contamination is unlikely. We did not estimate $\vartheta_{50}{}^{*}$ and $\vartheta_{95}{}^{*}$ because derived mean estimates of ϑ were uninformative.

3.2| Systemic Contamination

Naïve results : Amplification of contaminating DNA for all workflow steps was more common than in the sporadic contamination experiment (Table 2). Contamination was detected in all PCR replicates for the extraction-high contamination treatment (all 480 PCR replicates, mean ± standard deviation Cq = 33.85 ±0.81, and copy number estimate = 31.48 ±17.33), in all samples (group of 8 PCR replicates) and most PCR replicates for the PCR-high contamination treatment (319 PCR replicates, Cq = 37.66 ±0.80, copy number estimate = 1.4184514 ±0.79), and in most samples but few PCR replicates for the filtration-high contamination treatment (66 PCR replicates, Cq = 38.87 ±0.54, copy number estimate = 0.57 ±0.26). Contamination was rarely detected in samples or PCR replicates for workflow steps with low-level contamination (19 PCR replicates across all treatment levels, Cq = 38.39 ±0.59, copy number estimate = 0.79 ±0.29).

Modeled results . The most supported model was Ψ (Workflow step × level), ϑ(Workflow step × level), p (Workflow step × level). The WAIC value of this model was 230 WAIC units less than any other model. Posterior mean estimates of Ψ were ˜1.00 across all workflow step × contamination level combinations (Fig. 2b), indicating that contamination was detected in at least one sample for all combinations. Posterior mean estimates of ϑ were ˜ 1.00 for the filtration and extraction-high level contamination treatments, ˜ 0.70 for the filtration and extraction-low level contamination treatments, and ˜ 0.55 for the PCR high and low-level contamination treatments (Fig. 2b). Posterior mean estimates of p were 1.00 for the extraction-high level contamination treatment, ˜0.60 for PCR high- and low-level contamination, and near zero for all other treatment combinations (Fig. 2b).

Uncertainty in Ψ estimates was moderate (CI widths 0.29 – 0.63), with larger CI widths reflecting lower-level sample and PCR detection differences caused by the interaction of workflow step with contamination level. For example, Ψ estimates for the extraction workflow had the largest CI widths because extraction low and high-level estimates of ϑ and p had the biggest differences. Uncertainty in ϑ estimates was highest for filtration and extraction-low level contamination treatments because their respective p values were near zero; uncertainty was minimal for the PCR high and low-level treatments since contaminating DNA was detected in most PCR replicates. There was no uncertainty in the extraction-high level contamination estimate of ϑ because contaminating DNA was detected in all PCR replicates. Uncertainty for all p estimates was minimal. Post-hoc power analyses using mean parameters from the most supported model indicate that the CI widths for Ψ , ϑ, and p would not decrease with more PCR replicates; in fact, the point of diminishing returns was often 1 PCR replicate (Appendix 3).

The $\vartheta_{0.50}$* (i.e., the estimated number of water samples required to have 50% probability of detecting contamination) was one sample for all filtration and extraction-contamination level combinations. (Fig. 4) The $\vartheta_{0.95}$* was also one sample for filtration and extraction-high level contaminations, whereas three samples were needed to detect low level contamination at these workflow steps (Fig. 4). The $p_{0.50}$* and$p_{0.95}$* had greater variability across treatment level combinations. The $p_{0.50}$* and$p_{0.95}$* for filtration was 34 and 148 replicates for low contamination levels and 4 and 18 replicates for high contamination levels; for extraction

it was 17 and 75 for low contamination and 1 and 1 for high contamination; and for PCR it was 1 and 4 for low contamination and 1 and 3 for high contamination (Fig. 4).

4| Discussion

The ability to recognize positive eDNA detections as contamination continues to be an outstanding need in eDNA monitoring because the costs of false positives can be high (Jerde, 2021; Sepulveda, Nelson, et al., 2020). Most eDNA sampling protocols use negative controls associated with [?] 1 workflow step to alert to the potential for contaminated field samples, but there can be large variation in negative control sampling schemes (Sepulveda, Hutchins, et al., 2020). Here, we used empirical data and lab experiments to evaluate the power of negative controls to detect contamination. We found that a typical negative control scheme had minimal power to detect sporadic contamination but had very high power to detect systemic contamination. Below we discuss implications of these results to eDNA study design.

4.1| Sporadic contamination is a difficult issue to resolve

Sporadic contamination was nearly absent in a workflow that is commonly used by many eDNA monitoring programs. We only had 13 of 6180 PCR (0.2%) replicates amplify for contaminated DNA (Table 1), even though many samples and replicates were processed in parallel with a higher number of positives samples than is typical for a realistic sampling event. Other studies that have examined large numbers of negative controls also found zero or near-zero evidence of amplification. Smith and Goldberg (2020) analyzed 50 PCR negative control replicates, and Tingley, Coleman, Gecse, van Rooyen, and R. Weeks (2021) analyzed 44 PCR replicates; both studies reported no amplifications. Serrao, Reid, and Wilson (2018) analyzed 258 negative control samples and found that 98.4% had detections of less than 1 copy reaction^{-1}and Guillera-Arroita, Lahoz-Monfort, van Rooyen, Weeks, and Tingley (2017) analyzed 992 qPCR replicates and had 8 (0.8%) replicates amplify. The rare amplifications associated with the much higher sample numbers of our study, underscored by the apparent association between amplification rate and sample size in Table 1, and the high sample numbers in Guillera-Arroita et al. (2017) and Serrao et al. (2018) do suggest that the potential for sporadic contamination events increases with sample size owing to the very nature of small probabilities. Thus, detection of such low frequency events is likely to require high effort.

While these extremely low rates of sporadic contamination should be reassuring to eDNA practitioners and end-users, it is important to consider whether these amplifications are indeed indicative of contamination (i.e., detections of DNA targets in a reaction not attributable to the presence of those DNA targets in the sample matrix prior to sample collection) or are false-positive tests (i.e., detections of non-target DNAs caused by experimental or assay design flaws or PCR error). Bayes' Theorem reminds us to evaluate posterior probabilities in the context of base rates; when the base rate of contamination and the probability of amplification given contamination are low, then there is a considerable likelihood that these very few negative control amplifications were not caused by contaminating substances and were instead other sources of error ("false positive test"; Jerde, 2021). Presence of contaminated DNA in amplified negative controls can be sequence-verified, but this is costly and, as was the case in this study, can result in ambiguous results when DNA targets are short (<150 bp) and amplification occurs at late cycles (Crossley et al., 2020). A rerun of all negative control samples should provide additional insight as to the likelihood of contamination vs. false positives; though, our cumulative probability estimates (p^*) do suggest that there is still a low probability of consistently detecting contaminating DNA (Fig X). In addition, the assay's limit of detection (LOD) can be used as an informative threshold to identify when contamination is likely to influence the interpretation of field sample results. When negative control amplification rates are rare and their copy numbers are <LOD, then detection inferences from field samples with copy numbers > LOD should be minimally influenced. Alternatively, there is the potential for strong influence when negative control results are > LOD. There is value in the imperfect information provided by negative control amplification rates, base rates and limits of detection because they can be used to bracket confidence in eDNA results.

The non-zero probability for sporadic contamination and the near-zero probability of detecting sporadic contamination make accurate interpretation of eDNA results challenging, especially when non-molecular

methods cannot confirm the eDNA results. Multi-scale occupancy models that account for false positives at the sample and PCR levels may provide a means to bolster confidence in eDNA result interpretations. Even when there is no evidence of contamination (i.e., no negative controls amplified), low false-positive rates can be incorporated into occupancy models to evaluate how parameter estimates and inferences may change (Griffin, Matechou, Buxton, Bormpoudakis, & Griffiths, 2020; Smith & Goldberg, 2020). For instance, results interpretations that change with inclusion of non-zero false positive rates might reflect high uncertainty in the original results. In addition to statistical approaches, eDNA results should be interpreted in the context of the entire dataset. Rare amplification of negative controls that co-occur with rare amplification of field samples should be considered more suspect than no or rare amplification of negative controls that co-occur with frequent amplification of field samples. If systemic contamination occurred, then our results indicate that amplification of negative controls and field samples should both be frequent.

4.2| Negative control effort

If we assume that our sporadic negative controls amplified because of contamination, then our modeled estimates of p^*suggest that the required effort to consistently detect sporadic contamination is not tractable. A coin toss of detecting sporadic contamination (p_{50}*) requires an average (± 95% CI) of 106 (38 –287) PCR replicates per negative control, whereas a near guarantee (p_{95}*) requires an average of 458 (167-1241) PCR replicates per negative control (Figure 3). Comparison of these estimates to two prominent eDNA surveillance programs underscores how this level of effort is out of reach. The Asian carp (*Hypophthalmichthys s* pp.*)* eDNA surveillance program requires that a minimum of 10% of the number of samples collected should be field negative controls; these field negative controls are only analyzed at 8 PCR replicates, and only 4 NTCs are run per plate (US Fish and Wildlife Service, 2020). The great crested newt (*Triturus cristatus*) eDNA surveillance program only includes an extraction negative control, which is run with 12 PCR replicates and 4 NTCs per plate (Biggs et al., 2014). Comparable negative control schemes are common in ancient DNA analyses (e.g., Fulton and Stiller 2012) and DNA forensics (Moore & Kornfield, 2012; Parson et al., 2014). In comparison to the intractability of detecting sporadic contamination, detection of systemic contamination is an attainable objective for eDNA monitoring programs. Near-certain ($\vartheta_{0.95}$*) detection of low-level contamination in a sample required on average two negative field and extraction control samples and near-certain ($p_{0.95}$*) detection of low-level contamination in PCR negative controls required on average 4 PCR replicates.

4.3| Study design considerations

Our results have important implications for eDNA study design if the objective is to determine if field samples are compromised by systemic contamination, rather than at which workflow step potential contamination occurred. We could not reliably identify workflow steps where contamination occurred, but with lower levels of contamination anywhere in workflow, we could consistently detect systemic contamination. First, a high ratio of negative controls to field samples may not be an efficient use of resources because very few negative control samples are required to have high confidence that systemic contamination is absent; magnitudes more negative control samples are required for marginal confidence that sporadic contamination is absent. Second, increasing the number of PCR replicates per negative control sample rather than the number of negative control samples results in a higher probability of detecting systemic contamination. Our estimates of$p$ * for the filtration and extraction steps increased with the number of replicates, whereas estimates of ϑ* were invariant to the number of samples. Third, the number of negative control samples (and their PCR replicates) at the earliest stage of the eDNA workflow should be maximized relative to negative control samples associated with later stages. When early-stage negative controls are handled in the same manner as samples through all workflow stages, they can provide a comprehensive screening of systemic contamination (Goldberg et al., 2016). Other disciplines that use PCR-based methods to amplify samples with little target DNA have also identified that inclusion of negative controls at early stages of the workflow bolsters confidence in results (e.g., Weyrich et al. 2019).

Given that most eDNA negative control sampling schemes are only effective at detecting systemic contamination and not sporadic contamination, the nuances of a negative control may be less important than previously

thought (Sepulveda, Hutchins, et al., 2020). Nuances include collecting field negative controls once per site vs. once per day, as the first vs. last sample collected at a site, or laboratory (e.g., deionoized water) vs. environmental (e.g., presumed negative field site) water sources. Different negative control methods are likely to provide similar results when contamination is systemic (e.g., contamination of laboratory reagents).

5| Summary

Sample contamination in eDNA studies is a rare occurrence when best-practices are employed. As a result, tractable control schemes are inadequate to demark when sporadic contamination has occurred. A statistical modeling approach, wherein sporadic contamination rates are incorporated in detection models to reflect uncertainty, is likely the best way to manage sporadic contamination risks. In contrast, systemic contamination at high amounts was very reliably detected by a tractable number of negative controls (between one and three replicates) at each eDNA workflow step. The effectiveness of these samples to detect contamination was, however, greatly diminished when the concentration of introduced contamination was one to two orders of magnitude below the high-level contamination. Due to the hierarchical nature of the eDNA workflow, current control schemes are inadequate to source-trace contamination to a particular part of the workflow. Therefore, an efficient approach would include a greater proportion of negative controls introduced in the earliest workflow steps so that these negative controls survey the entire workflow for contamination.

5| Acknowledgements

We thank K. Klymus (USGS) and Yale Passamaneck (U.S. Bureau of Reclamation) for providing initial reviews of this manuscript. This study was funded by the USGS Mid-continent Region.

6| Author Contributions

- designed research (PRH, AJS)
- performed research (PRH, LNS, AJS)
- analyzed data (PRH, AJS)
- wrote the manuscript (PRH, AJS)
- manuscript editing and revision (PRH, LNS, AJS)

7| Data accessibility and Benefit-Sharing Statement

eDNA technical replicate and sample results are accessible from the USGS Sciencebase at doi: XXXXXXXXXX

Figures and Tables

Table 1. Proportion (Prop) of negative control PCR replicates and samples that amplified in sporadic contamination assessments (i.e., when contamination was not intentionally introduced but negative samples are processed in isolation or in parallel with highly positive samples).

		PCR	PCR	Sample	Sample
Workflow step	**Parallel processing**	**Prop**	**n**	**Prop**	**n**
Field	All negative	0.003	960	0.025	120
Filtration	All negative	0.002	960	0.017	120
	Mixed	0.000	480	0.000	60
Extraction	All negative	0.002	960	0.017	120
	Mixed	0.000	480	0.000	60
PCR	All negative	0.005	960	–	–
	Mixed	0.002	480	–	–
	Standard curve	0.000	900	–	–

Table 2. Proportion positive samples (n = 60) and positive PCR replicates (n =480) for systemic contam-

ination raw results.

Workflow step	Contamination level	PCR	Sample
Filtration	Low	0.01	0.08
	High	0.14	0.65
Extraction	Low	0.02	0.13
	High	1.00	1.00
PCR	Low	0.01	–
	High	0.66	–

Figures

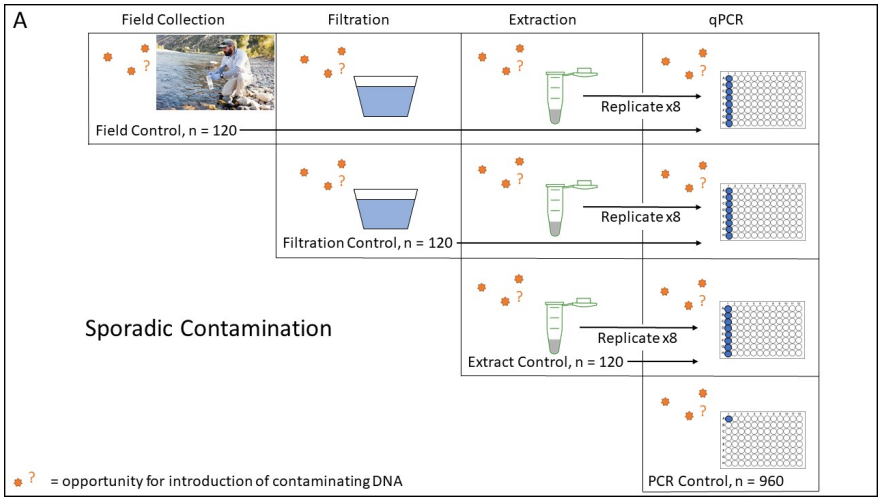

Figure 1. A schematic showing the experimental design for sporadic contamination (A) and systemic contamination (B). Labels at the top of each diagram demark the workflow step where the corresponding samples were initiated. Labels inside or next to boxes indicate the treatment level and the total number of replicates performed at the initial workflow step.

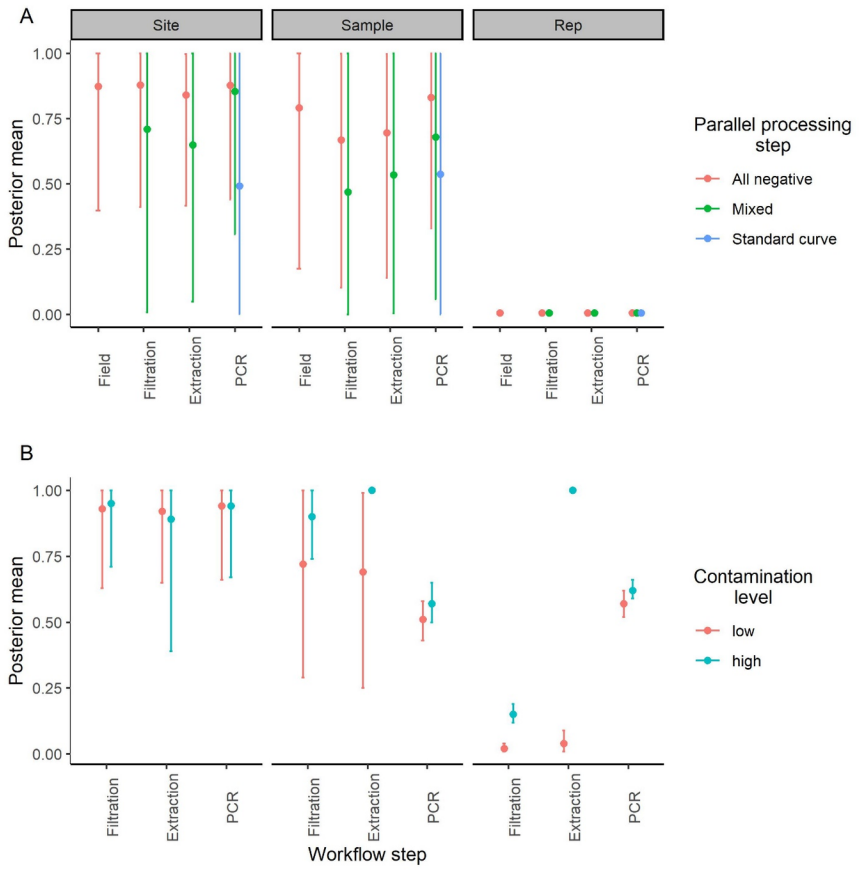

Figure 2. Posterior mean probabilities (± 95% credibility intervals) of detecting sporadic (A) or systemic contamination (B) (y-axes) at the site, sample replicate, or PCR replicate level (panels) for each workflow step (x-axes) when contamination is not intentionally introduced but negative samples are processed in isolation (All negative) or in parallel with highly positive samples (Mixed) or six-point standard curve series (Standard curve) (A) and when contamination is intentionally introduced to every sample at either a high or low concentration (B).

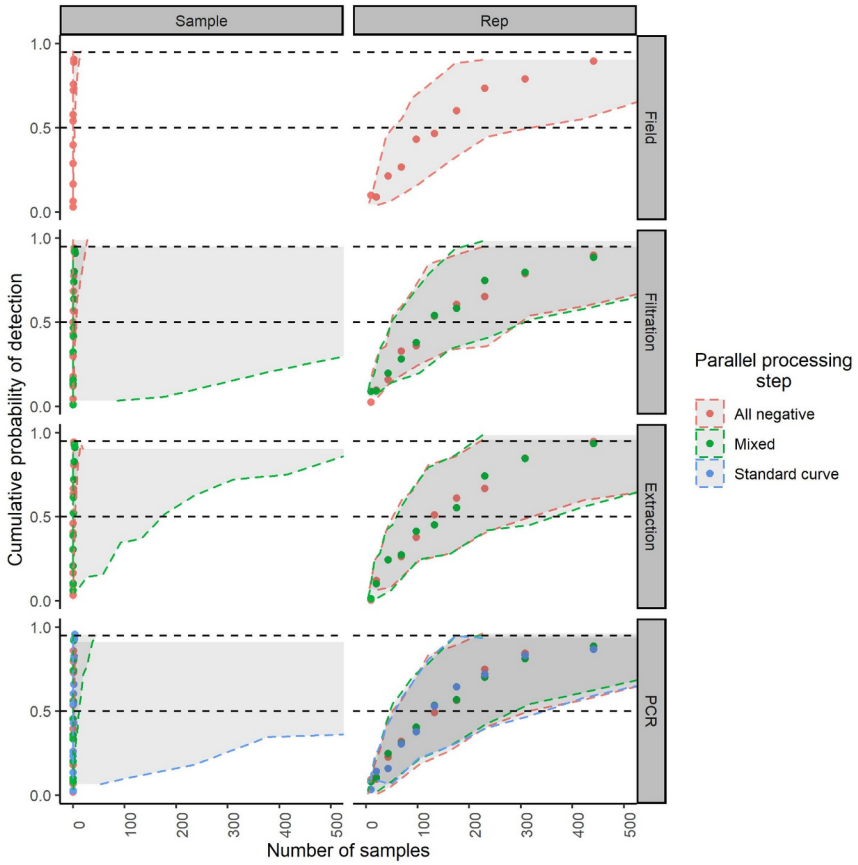

Figure 3. The cumulative probability of detecting sporadic contamination with different replicate effort at the sample replicate or PCR replicate level (x-panels) for each workflow step (y-panels) when contamination is not intentionally introduced but negative samples are processed in isolation (All negative) or in parallel with highly positive samples (Mixed) or six-point standard curve series (Standard curve). Shaded areas indicate 95% credibility intervals for each treatment level.

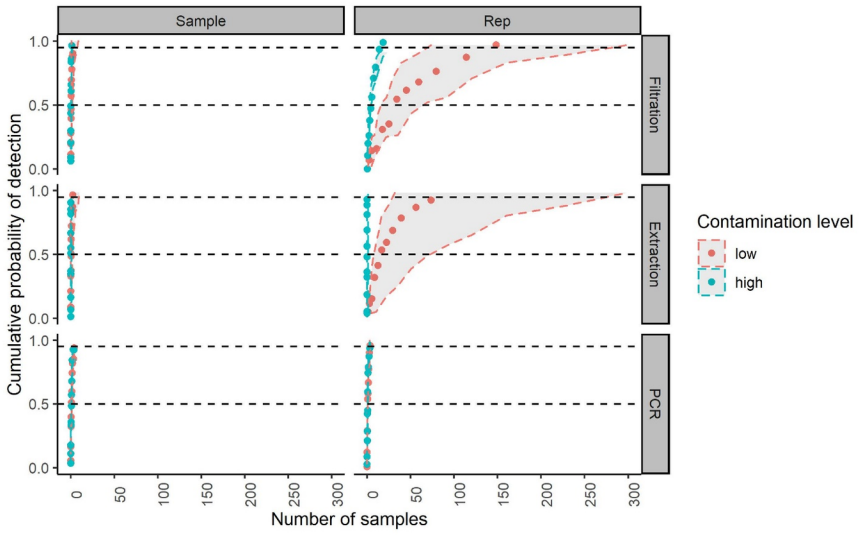

Figure 4. The cumulative probability of detecting systemic contamination with different replicate effort at the sample replicate or PCR replicate level (x-panels) for each workflow step (y-panels) when contamination is intentionally introduced to every sample at either a high or low concentration. Shaded areas indicate 95% credibility intervals for each treatment level.

Literature Cited

Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics, 17* (1), 1-17. doi:10.1007/s10592-015-0775-4

Bass, D., Stentiford, G. D., Littlewood, D. T. J., & Hartikainen, H. (2015). Diverse Applications of Environmental DNA Methods in Parasitology. *Trends in Parasitology, 31* (10), 499-513. doi:https://doi.org/10.1016/j.pt.2015.06.013

Biggs, J., Ewald, N., Valentini, A., Gaboriaud, C., Griffiths, R. A., Foster, J., . . . Dunn, F. (2014). Analytical and methodological development for improved surveillance of the Great Crested Newt.*Defra Project WC1067* .

Borst, A., Box, A. T. A., & Fluit, A. C. (2004). False-Positive Results and Contamination in Nucleic Acid Amplification Assays: Suggestions for a Prevent and Destroy Strategy. *European Journal of Clinical Microbiology and Infectious Diseases, 23* (4), 289-299. doi:10.1007/s10096-004-1100-1

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., . . . Wittwer, C. T. (2009). The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments.*Clinical Chemistry, 55* (4), 611-622. doi:10.1373/clinchem.2008.112797

Carim, K. J., Padgett-Stewart, T. M., Wilcox, T. M., Young, M. K., McKelvey, K. S., & Schwartz, M. K. (2015). *Protocol for collecting eDNA samples from streams [Version 2.3]* . Retrieved from

Cristescu, M. E., & Hebert, P. D. N. (2018). Uses and Misuses of Environmental DNA in Biodiversity Science and Conservation. *Annual Review of Ecology, Evolution, and Systematics, 49* (1), 209-230. doi:10.1146/annurev-ecolsys-110617-062306

Crossley, B. M., Bai, J., Glaser, A., Maes, R., Porter, E., Killian, M. L., . . . Toohey-Kurth, K. (2020). Guidelines for Sanger sequencing and molecular assay monitoring. *Journal of Veterinary Diagnostic Investigation, 32* (6), 767-775. doi:10.1177/1040638720905833

Darling, J. A., Jerde, C. L., & Sepulveda, A. J. (2021). What do you mean by false positive? *Environmental DNA, n/a* (n/a). doi:https://doi.org/10.1002/edn3.194

Erickson, R. A., Merkes, C. M., & Mize, E. L. (2019). Sampling Designs for Landscape-level eDNA Monitoring Programs. *Integrated Environmental Assessment and Management, 15* (5), 760-771. doi:https://doi.org/10.1002/ieam.4155

Forootan, A., Sjöback, R., Björkman, J., Sjögreen, B., Linz, L., & Kubista, M. (2017). Methods to determine limit of detection and limit of quantification in quantitative real-time PCR (qPCR). *Biomolecular Detection and Quantification, 12* , 1-6. doi:https://doi.org/10.1016/j.bdq.2017.04.001

Fulton, T. L., & Stiller, M. (2012). PCR Amplification, Cloning, and Sequencing of Ancient DNA. In B. Shapiro & M. Hofreiter (Eds.),*Ancient DNA: Methods and Protocols* (pp. 111-119). Totowa, NJ: Humana Press.

Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., . . . Taberlet, P. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution, 7* (11), 1299-1307. doi:https://doi.org/10.1111/2041-210X.12595

Griffin, J. E., Matechou, E., Buxton, A. S., Bormpoudakis, D., & Griffiths, R. A. (2020). Modelling environmental DNA data; Bayesian variable selection accounting for false positive and false negative errors. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 69* (2), 377-392. doi:https://doi.org/10.1111/rssc.12390

Guillera-Arroita, G., Lahoz-Monfort, J. J., van Rooyen, A. R., Weeks, A. R., & Tingley, R. (2017). Dealing with false-positive and false-negative errors about species occurrence at multiple levels.*Methods in Ecology and Evolution, 8* (9), 1081-1091. doi:https://doi.org/10.1111/2041-210X.12743

Hutchins, P. R., Sepulveda, A. J., Hartikainen, H., Staigmiller, K. D., Opitz, S. T., Yamamoto, R. M., . . . Okamura, B. (2021). Exploration of the 2016 Yellowstone River fish kill and proliferative kidney disease in wild fish populations. *Ecosphere, 12* (3), e03436. doi:https://doi.org/10.1002/ecs2.3436

Jerde, C. L. (2021). Can we manage fisheries with the inherent uncertainty from eDNA? *Journal of Fish Biology, 98* (2), 341-353. doi:https://doi.org/10.1111/jfb.14218

Laramie, M. B., Pilliod, D. S., Goldberg, C. S., & Strickler, K. M. (2015). *Environmental DNA sampling protocol - filtering water to capture DNA from aquatic organisms* (2-A13). Retrieved from Reston, VA: http://pubs.er.usgs.gov/publication/tm2A13

Minamoto, T., Miya, M., Sado, T., Seino, S., Doi, H., Kondoh, M., . . . Uchii, K. (2021). An illustrated manual for environmental DNA research: Water sampling guidelines and experimental protocols.*Environmental DNA, 3* (1), 8-13. doi:https://doi.org/10.1002/edn3.121

Moore, M. K., & Kornfield, I. L. (2012). Best Practices in Wildlife Forensic DNA. *Wildlife Forensics* , 201-236. doi:https://doi.org/10.1002/9781119953142.ch11

Parson, W., Gusmão, L., Hares, D. R., Irwin, J. A., Mayr, W. R., Morling, N., . . . Parsons, T. J. (2014). DNA Commission of the International Society for Forensic Genetics: Revised and extended guidelines for mitochondrial DNA typing. *Forensic Science International: Genetics, 13* , 134-142. doi:10.1016/j.fsigen.2014.07.010

Patrone, P. N., Romsos, E. L., Cleveland, M. H., Vallone, P. M., & Kearsley, A. J. (2020). Affine analysis for quantitative PCR measurements. *Analytical and Bioanalytical Chemistry, 412* (28), 7977-7988. doi:10.1007/s00216-020-02930-z

R-Core-Team. (2014). R: A Language and Environment for Statistical Computing: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Sepulveda, A. J., Amberg, J. J., & Hanson, E. (2019). Using environmental DNA to extend the window of early detection for dreissenid mussels. *Management of Biological Invasions, 10* (2), 342-342–358.

Sepulveda, A. J., Hutchins, P. R., Forstchen, M., McKeefry, M. N., & Swigris, A. M. (2020). The Elephant in the Lab (and Field): Contamination in Aquatic Environmental DNA Studies. *Frontiers in Ecology and Evolution, 8* , 440.

Sepulveda, A. J., Nelson, N. M., Jerde, C. L., & Luikart, G. (2020). Are Environmental DNA Methods Ready for Aquatic Invasive Species Management? *Trends in Ecology & Evolution, 35* (8), 668-678. doi:https://doi.org/10.1016/j.tree.2020.03.011

Serrao, N. R., Reid, S. M., & Wilson, C. C. (2018). Establishing detection thresholds for environmental DNA using receiver operator characteristic (ROC) curves. *Conservation Genetics Resources, 10* (3), 555-562. doi:10.1007/s12686-017-0817-y

Smith, M. M., & Goldberg, C. S. (2020). Occupancy in dynamic systems: accounting for multiple scales and false positives using environmental DNA to inform monitoring. *Ecography, 43* (3), 376-386. doi:https://doi.org/10.1111/ecog.04743

Stratton, C., Sepulveda, A. J., & Hoegh, A. (2020). msocc: Fit and analyse computationally efficient multi-scale occupancy models in r.*Methods in Ecology and Evolution, 11* (9), 1113-1120. doi:https://doi.org/10.1111/2041-210X.13442

Tingley, R., Coleman, R., Gecse, N., van Rooyen, A., & R. Weeks, A. (2021). Accounting for false positive detections in occupancy studies based on environmental DNA: A case study of a threatened freshwater fish (Galaxiella pusilla). *Environmental DNA, 3* (2), 388-397. doi:https://doi.org/10.1002/edn3.124

US Fish and Wildlife Service. (2020). *Quality Assurance Project Plan: eDNA Monitoring of Bighead and Silver Carps* . Retrieved from Bloomington, MN:

Weyrich, L. S., Farrer, A. G., Eisenhofer, R., Arriola, L. A., Young, J., Selway, C. A., . . . Cooper, A. (2019). Laboratory contamination over time during low-biomass sample analysis. *Molecular Ecology Resources, 19* (4), 982-996. doi:https://doi.org/10.1111/1755-0998.13011