VTRC
Virginia Transportation Research Council

We Bring Innovation to Transportation

Refining Environmental DNA Protocols Developed for the Endangered James Spinymussel (*Pleurobema collina*)

http://www.virginiadot.org/vtrc/main/online_reports/pdf/21-r14.pdf

RODNEY J. DYER
Professor and Director

BONNIE A. RODERIQUE
Field and Molecular Technician

M. CHARIS DEADWYLER
GIS Technician

Center for Environmental Studies
Virginia Commonwealth University

Final Report VTRC 21-R14

VIRGINIA TRANSPORTATION RESEARCH COUNCIL
530 Edgemont Road, Charlottesville, VA 22903-2454
vtrc.virginiadot.org

Standard Title Page - Report on Federally Funded Project

1. Report No.: FHWA/VTRC 21-R14	2. Government Accession No.:	3. Recipient's Catalog No.:
4. Title and Subtitle: Refining Environmental DNA Protocols Developed for the Endangered James Spinymussel *(Pleurobema collina)*		5. Report Date: February 2021 6. Performing Organization Code:
7. Author(s): Rodney J. Dyer, Bonnie A. Roderique, and M. Charis Deadwyler		8. Performing Organization Report No.: VTRC 21-R14
9. Performing Organization and Address: Virginia Transportation Research Council 530 Edgemont Road Charlottesville, VA 22903		10. Work Unit No. (TRAIS): 11. Contract or Grant No.: 114306
12. Sponsoring Agencies' Name and Address: Virginia Department of Transportation 1401 E. Broad Street Richmond, VA 23219 Federal Highway Administration 400 North 8th Street, Room 750 Richmond, VA 23219-4825		13. Type of Report and Period Covered: Final Contract 14. Sponsoring Agency Code:
15. Supplementary Notes: This is an SPR-B report.		

16. Abstract:

Molecular genetic techniques provide tools that may be used to locate, monitor, and survey the presence of cryptic aquatic species such as the endangered James Spinymussel (*Pleurobema collina*). Previous work on this species developed protocols that—across a range of conditions—were at least as accurate for detecting the presence of *P. collina* as currently deployed physical sampling approaches. This project builds on those findings in three different ways. First, the limits of detection for realistic field conditions were examined and showed that for native populations whose census sizes are on the order of 10 to 20 known individuals, environmental (eDNA) approaches can similarly provide positive evidence of *P. collina* presence. Second, alternative approaches for both DNA extraction and post-extraction cleanup were tested to identify the most effective combination that reduces the negative effects of inhibition on successful PCR amplification. Tests revealed two viable options that recover up to 56% of samples previously yielding no PCR product due to environmental inhibition. Third, *in situ* transect sampling was performed at two different locations to evaluate the extent to which eDNA template concentration can be used to model downstream concentration gradients. Across two sites and repeated sampling sessions, we were not able to construct well supported diffusion approximations for eDNA template concentrations for distances extending up to 500 meters downstream of known populations. The results of this work show that the eDNA approach for *P. collina* surveys is a viable addition to the suite of tools available to agencies for monitoring and management of this cryptic aquatic species. The study recommends that the Virginia Department of Transportation's Environmental Division meet with the U.S. Fish and Wildlife Service and the Virginia Department of Wildlife Resources to discuss whether there are circumstances under which eDNA can be an acceptable replacement for in-person surveys.

17 Key Words: Environmental DNA, eDNA, James Spineymussel, molecular genetics	18. Distribution Statement: No restrictions. This document is available to the public through NTIS, Springfield, VA 22161.

19. Security Classif. (of this report): Unclassified	20. Security Classif. (of this page): Unclassified	21. No. of Pages: 22	22. Price:

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

FINAL REPORT

REFINING ENVIRONMENTAL DNA PROTOCOLS DEVELOPED FOR THE ENDANGERED JAMES SPINYMUSSEL (*Pleurobema collina*)

Rodney J. Dyer
Professor and Director

Bonnie A. Roderique
Field and Molecular Technician

M. Charis Deadwyler
GIS Technician

Center for Environmental Studies
Virginia Commonwealth University

VTRC Project Manager
Bridget M. Donaldson, Virginia Transportation Research Council

In Cooperation with the U.S. Department of Transportation
Federal Highway Administration

Virginia Transportation Research Council
(A partnership of the Virginia Department of Transportation
and the University of Virginia since 1948)

Charlottesville, Virginia

February 2021
VTRC 21-R14

DISCLAIMER

The project that is the subject of this report was conducted under contract with the Virginia Department of Transportation, Virginia Transportation Research Council. The contents of this report reflect only the view of the authors, who maintain responsibility for the facts and accuracy of the data presented herein. The contents of this document do not reflect the views or policies of the Virginia Department of Transportation, the Commonwealth Transportation Board, or the Federal Highway Administration. This report does not claim to constitute a standard, specification, or regulation. Use of manufacturer names, trade names, or trademarks herein are solely for identification purposes and are not to be considered as endorsements.

Each contract report is peer reviewed and accepted for publication by staff of the Virginia Transportation Research Council with expertise in related technical areas. Final editing and proofreading of the report are performed by the contractor.

ABSTRACT

Molecular genetic techniques provide tools that may be used to locate, monitor, and survey the presence of cryptic aquatic species such as the endangered James Spinymussel (*Pleurobema collina*). Previous work on this species developed protocols that - across a range of conditions - were at least as accurate for detecting the presence of *P. collina* as currently deployed physical sampling approaches. This project builds upon those findings in three different ways. First, we examined the limits of detection for realistic field conditions and showed that for native populations whose census sizes are on the order of 10 to 20 known individuals, environmental (eDNA) approaches can similarly provide positive evidence of *P. collina* presence in each case. Second, we tested alternative approaches for both DNA extraction and post-extraction cleanup to identify the most effective combination that reduces the negative effects of PCR inhibition. Tests revealed two viable options that recovered up to 56% of samples previously yielding no PCR product due to environmental inhibition. Third, we performed *in situ* transect sampling at two different locations to evaluate the extent to which eDNA template concentration can be used to model downstream concentration gradients. Across two sites and repeated sampling sessions, we were not able to construct well supported diffusion approximations for eDNA template concentrations for distances extending up to 500 meters downstream of known populations. The results of this work show that the eDNA approach for *P. collina* surveys is a viable addition to the suite of tools available to agencies for monitoring and management of this cryptic aquatic species.

FINAL REPORT

REFINING ENVIRONMENTAL DNA PROTOCOLS DEVELOPED FOR THE ENDANGERED JAMES SPINYMUSSEL (*Pleurobema collina*)

Rodney J. Dyer
Professor and Director

Bonnie A. Roderique
Field & Molecular Technician

M. Charis Deadwyler
GIS Technician

Center for Environmental Studies
Virginia Commonwealth University

INTRODUCTION

Choosing the appropriate methodology for identifying the presence and abundance of species is difficult if the taxon naturally occurs at low densities or is cryptic in nature. For freshwater mussels, physical survey approaches are commonly challenged by small, isolated populations within restricted geographic ranges (e.g., Strayer *et al.* 1996, Smith *et al.* 2001). Despite these challenges, appropriate sampling protocols are necessary to identify the presence of cryptic species, particularly in the area of proposed or ongoing project impact. If the species is also endangered, the Virginia Department of Transportation (VDOT) faces additional regulatory challenges in both planning and execution of maintenance and development programs for existing and new infrastructure. As a result, field sampling approaches should be applied that maximize the probability of correctly identifying the presence of these cryptic taxa while at the same time allowing for a probabilistic estimation of sampling error rates (e.g., not detecting the taxon even though it is present - a Type II statistical error). This work builds upon a set of molecular genetic techniques developed by Dyer and Roderique (2017) for identifying the presence and abundance of the James Spinymussel (*Pleurobema collina*), an endangered freshwater mussel.

Pleurobema collina (*P. collina*) is a freshwater mussel endemic to the lotic aquatic habitats in non-tidal streams of the James River basin of Virginia (Hove & Neves 1994). Ongoing landscape modification, sedimentation from upland sources, and competition from invasive species such as the Asian Clam (*Corbicula fluminea*) have increased habitat fragmentation (Petty 2005). While Hove (1990) reported that seven common fishes may serve as hosts for the Spinymussel, the abundance of these Cyprinidae (minnow) species has not changed dramatically suggesting that the decline is not due to host availability (Petty 2005). Over the last couple of decades, the cumulative effects of these impacts have resulted in a loss of over 90% of the species native range, resulting in the species being listed under the Endangered Species Act.

At present, the Virginia Department of Wildlife Resources (VDWR) has implemented an extensive conservation program designed to locate populations and conserve critical habitat.

Environmental DNA (eDNA) is a non-invasive means of detecting the presence of rare, endangered, or invasive species by isolating discrete pieces of both nuclear (nDNA) and mitochondrial DNA (mtDNA) from the water column. Minute particles of tissue, either excreted or shed from individuals *in situ*, is used as templates for DNA extraction and subsequent amplification using species specific genetic markers. Over the past two decades, an international coalition of researchers has developed targeted DNA sequences with sufficient diversity to be used as species-specific markers. These "DNA Barcodes" have been used to classify taxa ranging from nematodes to elephants (Eggert *et al*. 2002; Floyd *et al*. 2002) and have recently been applied to both species identification and monitoring of presence/absence in aquatic habitats. Example applications include the identification of invasive species such as the silver carp in the Mississippi drainage (Hickcox 2011) and the American Bullfrog across Spain (Ficetola *et al.* 2008) as well as identifying the presence of cryptic species such as the Rocky Mountain Tailed Frog and the Idaho Giant Salamander (Goldberg *et al.* 2011).

In a previous study, Dyer and Roderique (2017) developed a set of genetic markers specifically designed to detect the presence of *P. collina* in both captive and native populations. However, two of the surveyed natural populations, whose presence had been historically verified by physical sampling, did not yield positive detection using molecular genetic techniques. The census size of these two sites were thought to be in the range of 4 to 5 censused individuals. Given the inability of eDNA targets at these two populations, there was concern about the lower level of detection for molecular genetic techniques for locations of small census size. In addition to the sensitivity to small size, it was noted that several samples had contaminants that inhibited the application of molecular genetic approaches, essentially increasing the rate of false negative ascertainment. While a protocol was developed for identifying when inhibition prevents mechanisms within the Polymerase Chain Reaction (PCR) workflow from being able to resolve target presence and absence—so that researchers could distinguish between inferences of no target species present versus not being able to determine if the species is present—the authors had recommended that further work focus on methods that can reduce the effects of these unknown inhibitors on collected samples.

In addition to these challenges, the authors noted that a positive identification of an eDNA target sequence in the water column at a specific locale does not necessarily mean that the population of individuals is immediately proximate. The amplification of target DNA from the water column directly requires that tissue, gametes, and other cellular material is motile and flowing with the water. As such, the authors hypothesized that the suspension of materials would systematically decrease as a function of distance from the source population. If these diffusion model approximations could be successfully modeled, the spatial location of the up-stream population may be easier to identify. The work presented herein builds upon that original research to help elucidate potential benefits and challenges to the use of eDNA as a valuable tool in cryptic species identity and management.

PURPOSE AND SCOPE

The purpose of this work was to provide additional baseline performance information on eDNA methods to allow agencies to determine if these approaches may fit into the larger toolset for surveying *P. collina*. While the previous work on this species by Dyer and Roderique (2017) showed itself to be applicable and deployable, the authors identified some notable limitations to the approach that required subsequent investigation. Those limitations are addressed herein.

Given the recognized challenges for broad use of eDNA as a survey techniques, the scope of this work included the following three objectives:

Objective 1: Limits of Detection - In the previous work, there was a large gap in the census sizes of populations surveyed. At the lower end were two populations whose census size was estimated as 4 to 5 individuals, whereas the next larger population had hundreds. While we were unable to gain positive detection in the smallest of populations, the lower limits of eDNA detection probabilities were further examined in the field by surveying additional known populations with small census sizes.

Objective 2: Environmental Inhibition - Organic and inorganic compounds in environmental samples may act to inhibit the mechanisms necessary for successful PCR amplification of environmental samples by binding either the template DNA or interfering with the ability of the polymerase enzyme to replicate DNA. While we can identify the presence of inhibition due to these compounds in samples yielding negative results, additional methods needed to be surveyed to determine which combination of DNA extraction and post-extraction cleanup protocols minimized the likelihood of environmental inhibition.

Objective 3: Distance Decay - Detecting a positive signal for cryptic species in freshwater streams only means that the population from which the target DNA template was derived is somewhere upstream of the sampling location. As the target for all eDNA amplifications is derived from individual animals, finding a small amount of template DNA at a specific location could mean that there is a small population in close proximity and the sample has sequestered a large fraction of the available template from that population, or it could mean that the template may be from a larger population further upstream and the remaining template derived from that population has already settled out from the water column. In this objective we used transect sampling approaches to determine if changes in the sampling design for eDNA studies can help develop additional insights by modeling a decay curve based upon measured target DNA concentration to better understand the spatial proximity of the population from which the template is being derived.

METHODS

Limits of Detection

Previous studies on this species focused on six sites where *P. collina* was part of an ongoing monitoring program and had been physically sampled several times. Individual census sizes in these populations ranged from 4 to 1,125 individuals per site. These two smallest populations, with census sizes estimated to be 4 to 5 individuals, were the only sites that did not yield positive indication of *P. collina* DNA. Given the protected status of this species, we were not able to conduct manipulative experiments in the field to find out the lower limits of detection. We had intended to work with the Harrison Lake Fish Hatchery as a component of this Objective as they are growing populations of *P. collina* for reintroduction. However, their propagation program is raising individuals at levels of biomass and densities that are well above the lower regions range of eDNA detection. With the guidance of a VDWR biologist, we were able to identify three additional native populations (denoted as WC, RC, and SR) whose low census size allowed us to explore the lower limits of detection for *P. collina*.

Sampling proceeded as follows. The proposed size of the population at each site dictated the distribution of sampling. The physical distribution of individuals at RC was larger than those at the other two sites, so we increased the number of sampling locations at the site from three to five. At each location, we collected two replicate eDNA samples. Each sample consisted of an independent filtering of 2L of water onto a single extraction filter. All water samples were taken using the eDNA Sampler from Smith-Root (Vancouver WA) following manufacturers protocols. Replicate eDNA filters were also collected and stored following manufacturers recommendation. Template DNA was extracted from each individual filter using the DNEasy extraction kits following the manufacturer's protocol.

The presence of *P. collina* DNA template was evaluated by amplification of species-specific genetic markers designed and described in Dyer and Roderique (2017). Verification of *P. collina* presence was double checked by Sanger sequencing of amplified bands and comparing to known (both derived in this study and available from publications) genetic sequences of the target and potentially co-occurring taxa as reported previously. We report the results of these findings in terms of the fraction of sampled filters that yielded positive identification for *P. collina*.

Environmental Inhibition

Environmental inhibitors present in samples can impact the reaction efficiency of PCR by binding to nucleic acids, changing their chemical properties, or reducing the specificity of the primers (Abbaszadegan *et al*. 1993, John 1992, Opel *et al*. 2010). For water samples, the most likely inhibitors present are dissolved or solid organic compounds such as fulmic acids, humic acids, humic material, metal ions, and polyphenols (Abbaszadegan *et al*. 1993, Ijzerman *et al.* 1997). The impact of these inhibitors can be reduced either by diluting the sample (and hence the inhibitory agents) or identifying and removing the specific inhibitors directly.

To address this, we examined both DNA Extraction approaches as well as post-extraction cleanup techniques to determine if there were combinations of these that reduced inhibition the most. Samples for these tests were collected from field location identified in Dyer and Roderique (2017) that exhibited known PCR inhibition. Each sample was verified *a priori* to be exhibiting PCR inhibition prior to being used in this Objective. A water sample from each site was partitioned such that the same source locale would be subjected to both extraction protocols as well as all four post-cleanup approaches (Figure 1). The results will be presented as the fraction of samples that were inhibited, inconclusive, and yielding positive results.

Water samples known to contain inhibitory compounds that prevent PCR amplification were divided and each subjected to the following extraction techniques:

- *DNeasy + Qishredder (DQ):* The DNeasy is a silica-based DNA extraction approach commonly used in many applications of DNA extraction and amplification, including that from Dyer and Roderique (2017). To this we added the use of the Qiashredder is a unique biopolymer shredding system in a microcentrifuge spin-column that helps homogenize cells, tissue lysates, and other high molecular weight cellular components. It is chemically inert and will not bind to nucleic acids.

- *Powerwater (PW):* The Powerwater extraction isolates genomic DNA from water samples by precipitating salts, metals, humic substances, and other organic materials. This is a system specifically designed for isolating DNA from contaminated water

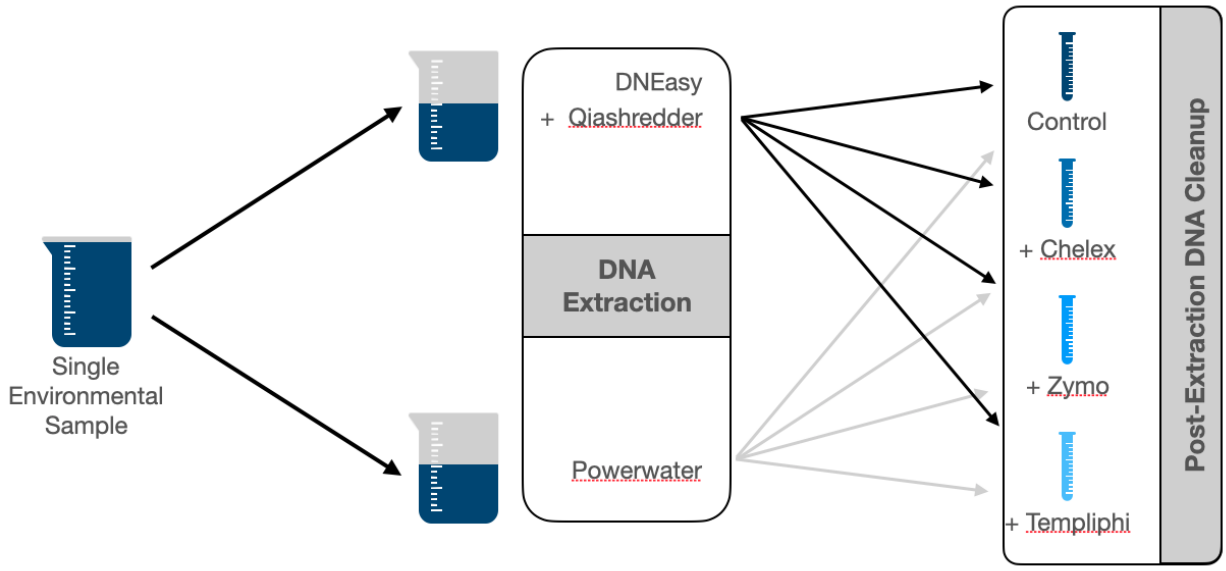

Figure 1. Sample allocation to test the effectiveness of alternative DNA extraction methods (DNeasy + Qiashredder vs Powerwater) and post-extraction cleanup (Control, Chelex, Zymo, and Templiphi). Each of 9 environmental samples were split and subjected to the complete 2 x 4 block design.

For both extraction techniques, replicates of extracted samples were then tested on each of the post-extraction cleanup protocols (following Hu *et al.* 2015). Different clean-up procedures target different sources of contamination. While we do not know the identity of

specific contaminants causing inhibition, the following set of protocols target broad classes of approaches for post-extraction cleanup.

- *Control*: No post-extraction cleanup was performed. Here the influences of alternative extraction methods are compared directly.

- *Chelex:* Chelex is a chelating resin developed for extracting DNA from samples for use with the PCR that has been shown in forensic science to be successful for samples of very small DNA concentrations with environmental contaminants.

- *Templiphi:* This is a novel, isothermal DNA amplification approach that uses the phi29 DNA polymerase enzyme and rolling circle amplification to generate high-quality templates for DNA for subsequent PCR.

- *Zymo:* This is an elution approach that clean extracted samples from DNA polymerases, modifying enzymes, RNA polymerases, ligases, kinases, nucleases, phosphatases, and restriction endonucleases.

Distance Decay

Detecting a positive signal for cryptic species, such as *P. collina*, in freshwater streams only means that the individuals are located upstream of the sampling location. Positive detection of an estimated concentration of template DNA could mean that there is a population in close proximity to the sampling location or a larger population further upstream whose remaining template has already settled out of the water column. In this objective we used transect sampling approaches to build a statistical representation of the DNA template concentration decay curve.

Sampling for this objective was replicated at two locales of known population size: DC/LOC (estimated census size exceeding 1500 individuals), and JC (estimated to have 428 individuals). For each location, complete sets of samples were collected on each of three separate sampling days. A single sampling transect and associated sampling locations were established on the first sampling session and water was collected at the same sites during each sampling trip. Transects were established starting at the downstream edge of the known population and extending 500 meters downstream. At marked 50 meter intervals along the transect, two independent water samples were collected and filtered using the eDNA backpack filter system (Smith-Root, Vancouver WA). Each sample was split and subjected to both the *DQ* and the *DQ_Zymo* cleanup procedure as described above as these were found to minimize the detrimental consequences of environmental inhibition (see Results Section). DNA concentration for each category of data at each sampling location was estimated using the protocols established in Dyer and Roderique (2017).

The systematic change in eDNA concentrations was examined using a hierarchical approach. At the highest level, all data at each locale was combined into a single model representing a location-wide estimate of how DNA concentrations decreased with distance from the source population. At a more fine-grained level, each sampling day was also fit to its own

model to capture any variation present between individual sampling days. The underlying distribution and error structure of the DNA concentration data was estimated using the Kolmogorov-Smirnov goodness of fit statistic and evaluated using Akaike's Information Criterion. Once the distribution of error terms were identified, individual decay models were fit using a generalized additive model for location scale and shape (GAMLSS; Rigby et al. 2019) with the distance downstream of the sampling location as the independent variable. Error terms and residuals will be examined for appropriateness following standard procedures. The *DQ* and *DQ + Zymo* data sets were analyzed separately to prevent pseudo-replication.

RESULTS

Limits of Detection

A VDWR biologist identified three potential sampling locations to evaluate the lower levels of detection for eDNA approaches. Physical census of these sites had estimated the census population sizes ranging from 10 to 20 individuals at each locale (Table 1; B Watson, personal communication). Each site was sampled once between July 30 to 8 August 2019 and mean discharge rates for each site were recorded from the closest USGS monitoring station as 12.7 m3/s for both WC & RC 40.0 m3/s at SR. At each site, replicate sampling locations (Table 1) were established at the location of the known population. At each sampling location within a site, two independent samples were taken and eDNA template was filtered onto individual extraction filters. The presence of *P. collina* eDNA template was identified at all of the three sampling locations (Table 1). The fraction of replicate filters sampled that yielded a positive identification of *P. collina* DNA sequences ranged from 17%-80% across sites.

Table 1. Sampling and detection probability for three populations of small estimated census size. At each site, a pair of eDNA filters were collected at several locations. Site names denote individual stream reaches. Exact site names are not reported given this species' status as endangered. The presence of *P. collina* DNA is indicated as well as the fraction of individual sample filters (denoted Filter Positive) yielding positive *P. collina* DNA.

Site	*Census*	*Locations*	eDNA Present	Filter Positive
WC	10-15	3	Yes	33%
RC	20	5	Yes	20%
SW	20	3	Yes	17%

Environmental Inhibition

Samples for this objective were collected from four different native populations (JC, LOC, MC, & DC) found by Dyer and Roderique (2017) to have high levels of PCR inhibition. PCR inhibition in these samples were verified prior to use in this test. For each of the 8 combinations of individual extraction methods (*DQ* & *PW*) and post-extraction cleanup (*Control*, *Chelex*, *Templiphi*, & *Zymo*), a total of N=9 replicate samples were tested and each sample was divided equally among both extraction and post-extraction treatments (as in Figure 1).

Independent of post-cleanup methodology, samples subjected to the DNeasy + Qishredder (*DQ*) extraction yielded samples with more concentrated DNA, a higher fraction of positive DNA profiles (average 44%) and lower fraction of verifiable inhibition (14%) than those using the *Powerwater* protocol (8% and 17% respectively; Table 2). For *DQ* samples, both the *Control* and the *Zymo* produced the highest fraction of positive results, whereas for *PW* there was little variation in positivity across the board that can be attributed to different post-extraction cleanup methods. Of note, while the fraction of samples recovered from inhibition is the same for these two approaches, the identity of these samples consist of a non-overlapping set—they do not recover the same set of samples, just the same overall fraction.

Distance Decay

The systematic change in eDNA concentrations was estimated using a transect sampling approach at two different locations: DC/LOC and JC. During each sampling period, stream water conditions were recorded (Table 3). When possible samples at both DC/LOC and JC were collected on the same day, though the final sampling session was done on successive days due to larger than expected flow rates at DC/LOC. There was no relationship between any measured stream condition and resulting eDNA detection probability.

Table 2. Estimated DNA concentrations for each treatment combination as well as the fraction of samples exhibiting inhibition (e.g., not recovered), and yielding positive *P. collina* DNA (e.g., inhibition removed and positive identification) using replicate extraction protocols: (A) DNeasy + Qiashredder (denoted as *DQ*) and (B) Powerwater (*PW*). For each combination of extraction and cleanup, the same set of 9 samples were subjected to all treatments.

A)

Extraction	Cleanup	Concentration	Inhibited	Positive
DQ	None	69.3	0	0.56
DQ	Chelex	12.1	0	0.22
DQ	Templiphi	578.5	0.22	0.44
DQ	Zymo	25.1	0.33	0.56

B)

Extraction	Cleanup	Concentration	Inhibited	Positive
PW	None	6.7	0	0.11
PW	Chelex	9.8	0	0.11
PW	Templiphi	490.0	0.67	0.11
PW	Zymo	4.6	0	0

Table 3. Stream conditions at individual sampling locations including acidity (pH), total dissolved solids, water temperature (°C), water conductivity, and flow (m^3/s) as well as resulting percentage of individual eDNA filters yielding a positive identification of *P. collina* DNA. Exact site names are not reported given this species' status as endangered.

Site	Date	pH	TDS	Temperature	Conductivity	Flow	eDNA Positivity
DC/LOC	2019-09-03	7.5	31.2	20.9	43.8	0.070	0.925
DC/LOC	2019-09-19	7.5	23.7	12.1	33.5	0.010	0.875
DC/LOC	2019-10-25	7.7	26.6	11.4	33.9	0.035	0.625
JC	2019-09-03	7.8	19.0	22.4	26.9	0.080	0.95
JC	2019-09-19	8.0	22.6	17.9	31.8	0.050	0.80
JC	2019-10-24	7.6	16.5	12.1	22.5	0.210	0.20

Quantitative PCR was run on N = 224 individual extraction filters, from which both *DQ* and *DQ+Zymo* extraction protocols and post-extraction treatment protocols yielded 448 samples (Figure A-1). Independent of sampling locale or time, samples extracted with *DQ* showed 1 sample (0.4%) exhibiting inhibition and 5 samples (2.2%) with no indication of *P. collina* DNA. The combination of *DQ* extraction and an additional *Zymo* cleanup yielded 17 inhibited samples (7.6%) and 8 (3.6%) with no indication of *P. collina* template DNA. Despite the differences in inhibition, the *DQ+Zymo* samples had significantly higher concentration of target DNA than the *DQ* extraction samples (t-test; t= 5.8598, df = 216, $P = 1.7e^{-8}$). As a result, samples from the *DQ+Zymo* extraction set were used to develop the following diffusion models due to increased concentrations. However, there was no difference in the interpretation or applicability of resulting models from the *DQ* data—data not shown for brevity.

Overall, mean *P. collina* eDNA concentration was both greater at JC than DC/LOC, and had more variability across sampling sessions (Table 4). As expected, most of the collected data was best described using a Gamma distribution, of which an exponential decay is a special case.

Table 4. Observed distribution of identified DNA concentration averaged across sampling sites and individually for each sampling day. Component refers to the sampling sites. Exact site names are not reported given this species' status as endangered

Component	Sample	Minimum	Mean	Maximum	Skew	Kurtosis	Distribution
DC/LOC	Combined	0.0001	1.6222	9.747	1.6277	8.1406	Gamma
DC/LOC	First	0.0390	1.9095	9.747	2.2139	10.5453	Log Normal
DC/LOC	Second	0.6970	2.4488	6.569	1.0704	6.3053	Log Normal
DC/LOC	Third	0.0001	0.0397	0.267	2.3026	8.8593	Gamma
JC	Combined	0.0001	3.4872	20.363	1.4964	5.4745	Gamma
JC	First	0.0001	3.8469	20.363	1.3325	4.0100	Gamma
JC	Second	0.0490	3.8945	9.166	0.2806	2.2753	Gamma
JC	Third	0.0001	0.1491	0.538	1.4170	2.9986	Gamma

At the site level, there did not appear to be a clear decay in concentrations across the 500 meter sampling transect (Figure 2). There does appear to be a trend in the percent positivity of individual eDNA filters as a function of distance at the JC location (Figure A-3), though this is only because on the third JC sampling trip high flow rates at the site resulted in failure to detect any positive samples beyond 250m downstream of the target population. The initial fits for models describing site-level patterns of eDNA concentration decay yielded mixed results (Table 5) with the decay parameter at JC being significant while decay parameters estimated at DC/LOC were not significantly different than zero. However, validation of the models showed the residuals from both locales to be deviant from model expectations making interpretation of either model questionable.

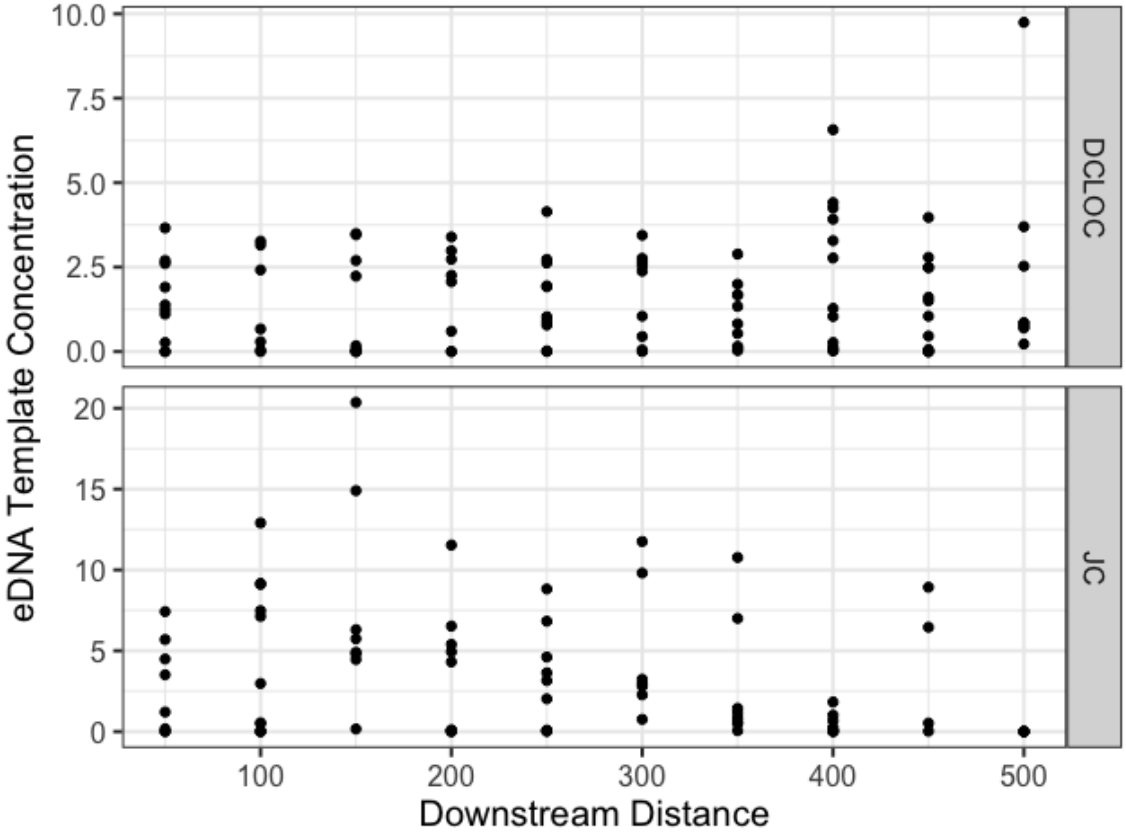

Figure 2. Estimated eDNA concentration as a function of distance downstream from known populations of *P. collina*. Data presented are from *DQ+Zymo* extraction. DC/LOC and JC denote sampling sites. Exact site names are not reported given this species' status as endangered.

Table 5. Parameters for distance decay models fit for each site. Parameters include the estimated decay parameter, degrees of freedom (df), the probability associated with keeping the parameter in the model (P_{Decay}) and the probability that the residual variation conformed to normality assumptions based upon Shapiro-Wilkes test ($P_{Shapiro}$). Exact site names are not reported given this species' status as endangered.

Site	df	Decay	P_{Decay}	$P_{Shapiro}$
DC/LOC	3,94	0.0008	0.4542	0.0000
JC	3,75	-0.0031	0.0482	0.0145

Models fit to individual sampling sessions were equally uninformative (Table 6). While overall models for the first and third sampling session at the DC/LOC location were significant, the sign of the decay function was positive, suggesting that DNA template concentrations actually increased with distance from the source population. For sampling sessions at JC, only the second sampling session had a decay function that was significantly different than zero; the residual variation, as quantified via a Shapiro-Wilks test, was also problematic. For all models, features other than simple downstream distance seems to be influencing estimated eDNA concentrations in non-linear fashion (see Figure A-2).

Table 6. Model parameters and summary statistics for individual decay regression models fit for each sampling session at both sites. Parameters include the decay parameter, the probability associated with keeping the decay parameter in the model (P_{Decay}) and the probability that the residual variation conformed to normality assumptions based upon Shapiro-Wilkes test ($P_{Shapiro}$). Exact site names are not reported given this species' status as endangered.

Site	Session	Decay	P_{Decay}	$P_{Shapiro}$
DC/LOC	First	0.0028	0.0299	0.0683
DC/LOC	Second	-0.0014	0.0212	0.7268
DC/LOC	Third	0.0115	0.0001	0.1297
JC	First	-0.0029	0.2671	0.0047
JC	Second	-0.0068	0.0000	0.0012
JC	Third	-0.0138	0.1284	0.1076

DISCUSSION

The results of this work show that molecular genetic approaches are viable additions to the suite of tools available to VDOT and VDWR biologists, particularly for use in monitoring and management of cryptic aquatic taxa. The results presented herein are in line with the previous work in terms of overall detection probabilities. At individual sites with a known resident population, we were able to make positive identification each time. Moreover, as the methodology requires multiple samples to be collected for each population, we found our per-filter detection probability to range from 17% to 33%, which is congruent with the observations from Dyer and Roderique (2017). If these are nascent detection probabilities for *P. collina* in these habitats, then future surveys benefit from estimation of necessary sampling effort required to gain a known likelihood of finding at least one positive filter using normal sampling intensity estimations. For these data and a conservative estimation of 17% detection, using 10 filters would yield an 85% probability of a positive result, 16 result in 95% probability, and 22 would yield a 99% detection probability.

We believe that the results presented here have been able to resolve some critical aspects that have been addressed in the previous work. First and foremost, due to the distribution of sampling locations used, Dyer and Roderique (2017) were not able to adequately determine if eDNA could identify the presence of the target species for the smallest populations (4 to 5 individuals). Unfortunately, the next larger population size in the original study had almost 175 individuals leaving a large gap of potential population sizes where molecular genetic approaches had not been identified. By expanding the sampling approaches here, to include additional

populations whose census size was estimated to be on the range of 10 to 20, this study was able to show positive identification for each site visited with multiple positive hits.

Another gap in our understanding had to do with inhibition of PCR processes by compounds found in the water samples. Evaluation of both extraction and post-extraction cleanup options showed a marked increase in the probability of recovering a positive template match over previous extraction approaches. While both *DQ* and *PW* were effective at removing inhibitors, it was surprising to find that PW also removed a disproportionately large fraction of the underlying template DNA as it was specifically designed to work on removing contaminants from water samples in preparation for PCR. It may be that the background template concentrations commonly found in eDNA analyses are too low to be effective and while *PW* precipitates free salts, metals, humic substances, and other organic materials, some template DNA is also removed bringing the overall concentration down below the limits of PCR effectiveness. In both laboratory and field testing, the *DQ* and the *DQ+Zymo* approaches had the highest probability of removing inhibitors, however, they removed inhibition from a partially non-overlapping set of samples. From a cost-effective perspective, the most efficient recovery of samples from inhibition are found when applying the DQ extraction approach and the DQ extraction and Zymo post-extraction cleanup methods simultaneously. The Zymo cleanup step only adds a few dollars per sample but the potential to rescue additional samples from inhibition is large, and as such we recommend simultaneous use of these techniques in situations where inhibition is a problem.

While these methods did show marked improvement over the previous extraction approach, it should be noted that the water samples for this study are largely clear and free of suspended particles. For systems with more stagnant water samples, the methods described above may not be as effective as reported herein. The design of each study should attempt to strike a balance between inhibitor removal and DNA template retention and will depend upon the purpose, scope, and budget of the particular study. And while we found success here, two points should be made. First, it should also be noted that we did not, nor did we intend to, identify the specific source or identity of the compounds causing inhibition; doing so would require samples to be processed by an environmental chemistry laboratory. Next, despite our findings of increased efficiency, there were still samples that showed inhibition, highlighting the need to make sure that any negative eDNA result be tested for inhibition to prevent potential false negatives.

Perhaps most surprising was the poor behavior of the decay functions for modeling the rate at which template DNA settles from the water column. While limited in applicability due the lack of robust statistical models derived to characterize the downstream settling rates of eDNA template, we have developed an online tool for managers interested in the habitat suitability of *P. collina*.

SUMMARY OF RESULTS

- *The lower limits of detection were verified for populations whose census size was as small as 10 to 20 individuals by expanding the sampling locations targeting isolated and diminutive populations.*

- *The effects of inhibition caused by materials and compounds within the water column can be mitigated to some degree by additional cleanup steps, though a balance between cleaning inhibitory compounds and reducing DNA template needs to be carefully considered.*

- *Across several sampling attempts, it was not possible to develop statistically suitable diffusion approximations for eDNA template settling velocities within a spatial distance of 500 meters downstream of known populations.* This suggests that positive eDNA results should be followed up by sampling considerably larger distances up-stream of the target location to be able to differentiate between a small proximate population and a larger population further upstream.

CONCLUSIONS

- *These results continue to suggest that eDNA may be a valuable sampling approach for DOTs. Decisions on its use to replace in-person surveys under certain circumstances should factor in habitat suitability and in-person sampling detection rates.*

- *The use of eDNA as an identification tool is widely practical and cost efficient.* However, it does not have the spatial specificity, at least within free flowing streams, to be a wholesale replacement of in-person sampling as highlighted by our inability to estimate appropriate decay models.

RECOMMENDATIONS

1. *VDOT's Environmental Division should meet with the U.S. Fish and Wildlife Service and the Virginia Department of Wildlife Resources to discuss whether there are circumstances under which eDNA is an acceptable replacement for in-person surveys.*

IMPLEMENTATION AND BENEFITS

Implementation

The implementation of Recommendation 1 will include staff from VDOT's Environmental Division, U.S. Fish and Wildlife Service, and the Virginia Department of Wildlife Resources. The findings from the study described herein and the initial study conducted by Dyer and Roderique (2017) can be used to inform decisions on the acceptable use of eDNA for *P. collina* surveys. VDOT's Environmental Division will contact the agencies by June 1, 2021 to schedule a meeting to discuss the implementation potential for the eDNA approach.

Benefits

The discussions that result from implementing Recommendation 1 will provide VDOT guidance with regard to circumstances under which eDNA may be an acceptable replacement for in-person surveys. Because *P. collina* can be difficult to survey given its cryptic nature, substantial VDOT resources are required in the form of labor and other associated in-person surveying costs. Outlining the conditions under which eDNA sampling can replace in-person surveys will ultimately increase the surveying efficiencies and decrease costs for VDOT. In addition, because the success of in-person sampling relies heavily on the expertise and diligence of the surveyor, the use of eDNA in some contexts may lower the risk of not detecting the taxon even if it is present.

ACKNOWLEDGMENTS

This work could not have been conducted without the assistance and guidance of Dr. Brian Watson, who provided raw tissue samples, location information, population demographics, stream access, and a wealth of knowledge.

REFERENCES

Abbaszadegan M, Huber MS, Gerba CP, Pepper IL, 1993. Detection of Enteroviruses in Groundwater with the Polymerase Chain Reaction. *Applied Environmental Microbiology*, 59, 5, 1318-1324.

Dyer RJ, Roderique BA. 2017. Development and testing of environmental DNA (eDNA) protocols for the endangered James Spinymussel (*Pleurobema collina*). Virginia Transportation Research Council. VTRC 18-R18.

Eggert LS, Rasner CA, Woodruff DS. 2002. The Evolution and Phylogeography of the African Elephant Inferred from Mitochondrial DNA Sequence and Nuclear Microsatellite Markers. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 269, 1504, 1993-2006.

Ficetola GF, Miaud C, Pompanon F, Taberlet P. 2008. Species Detection Using Environmental DNA From Water Samples. *Biology Letters*, 4, 4, 423-425.

Floyd R, Abebe E, Papert A, Blaxter M. 2002. Molecular Barcodes for Soil Nematode Identification. *Molecular Ecology,* 11, 4, 839-850.

Goldberg CS, Pilliod DS, Arkle RS, Waits LP. 2011. Molecular Detection of Vertebrates in Stream Water: A Demonstration Using Rocky Mountain Tailed Frogs and Idaho Giant Salamanders. *PLoS one*, e22746.

Hickcox T, Hsu FC, Wong R, Brito-Robinson T. 2011. *eDNA Surveillance of Asian Carp on the St. Croix and Mississippi Rivers*. Environmental DNA Solutions, Granger, IN.

Hove MC. 1990. *Distribution and Life History of the Endangered James Spinymussel, Pleurobema collina (Bivalvia: Unionidae).* M.S. Thesis. Virginia Polytechnic Institute and State University, Blacksburg.

Hove MC, Neves RJ. 1994. Life History of the Endangered James Spinymussel Pleurobema collina (Conrad, 1837; Mollusca: Unionidae). *American Malacological Bulletin*, 11, 1, 29-40.

Hu Q, Liu Y, Ui S, Huang D. 2015. A comparison of four methods for PCR inhibitor removal. *Forensic Science International*, 16, 94-97.

Ijzerman MM, Dahling DR, Fout GS. 1997. A Method to Remove Environmental Inhibitors Prior to the Detection of Waterborne Enteric Viruses by Reverse Transcription-Polymerase Chain Reaction. *Journal of Virological Methods*, 63, 1, 145-153.

John ME. 1992. An Efficient Method for Isolation of RNA and DNA From Plants Containing Polyphenolics. *Nucleic Acids Research*, 20, 9, 2381.

Opel KL, Chung D, McCord BR. 2010. A Study of PCR Inhibition Mechanisms Using Real Time PCR. *Journal of Forensic Science*, 55, 1, 25-33.

Petty MA. 2005. *Distribution, Genetic Characterization, and Life History of the James Spinymussel, Pleurobema collina (Bivalvia: Unionidae) in Virginia and North Carolina.* Doctoral Dissertation. Virginia Polytechnic Institute and State University, Blacksburg, VA.

Rigby RA, Stasinopoulos DM, Heller GZ, De Bastiani F. 2019. Distributions for modeling location, scale, and shape: *Using GAMLSS in R*, Chapman and Hall/CRC.

Smith DR, Villella RF, Lemarié DP. 2001. Survey Protocol for Assessment of Endangered Freshwater Mussels in the Allegheny River. *Journal of the American Benthological Society*, 20, 1, 118-132.

Strayer DL, Sprague S, Claypool S. 1996. A Range-Wide Assessment of Populations of *Alasmidonta heterodon*, an Endangered Freshwater Mussel (Bivalvia: Unionidae). *Journal of the North American Benthological Society*, 15, 3, 308-317.

APPENDIX

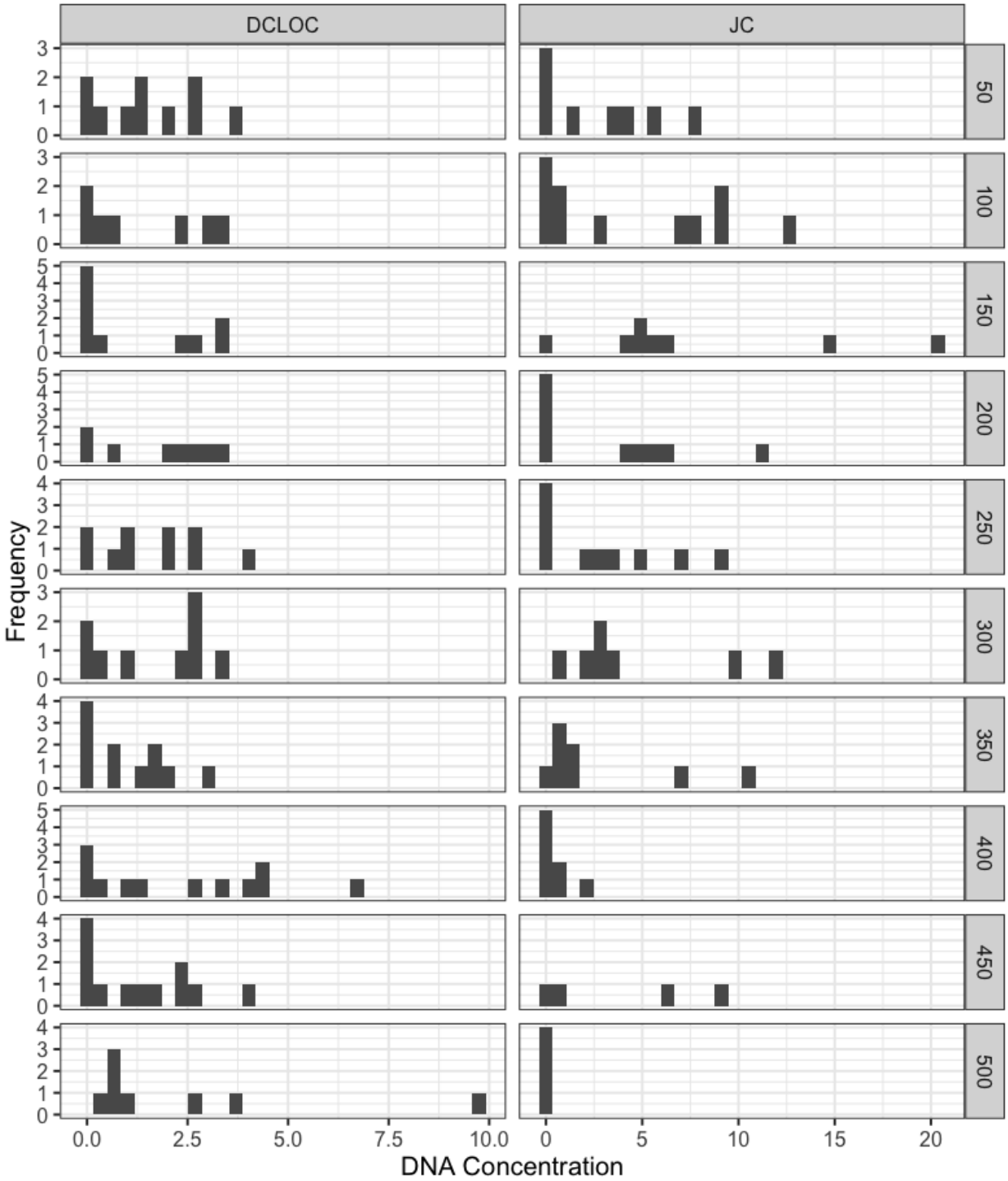

Figure A-1: The distribution of estimated eDNA concentrations for samples collected from three sampling sessions at DC/LOC and JC along sampling. DC/LOC and JC denote sampling sites. Exact site names are not reported given this species' status as endangered.

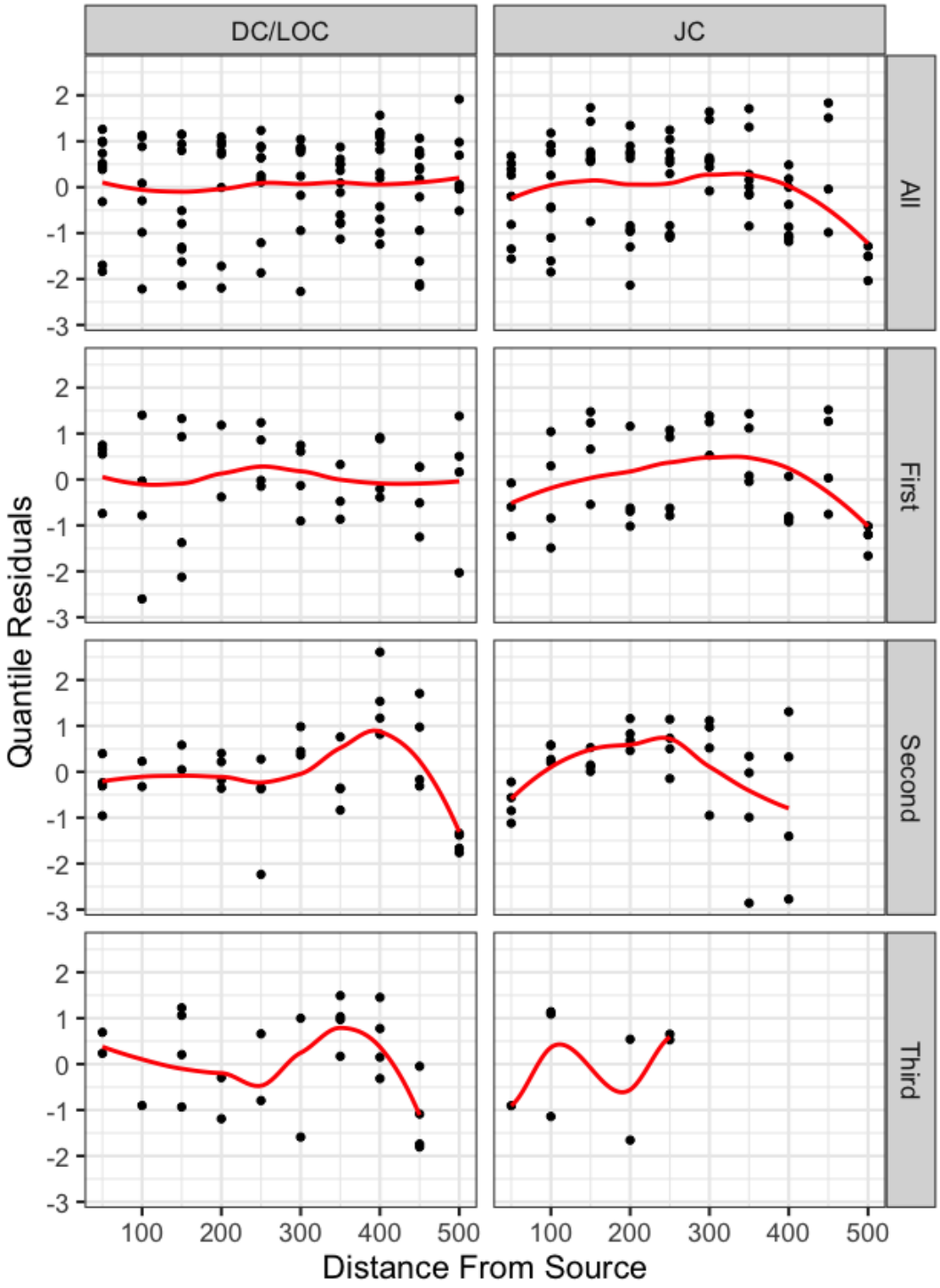

Figure A-2: Residual variation for decay models at both sites and across all sampling sessions. Red lines indicate statistical trend through the residuals. DC/LOC and JC denote sampling sites. Exact site names are not reported given this species' status as endangered.

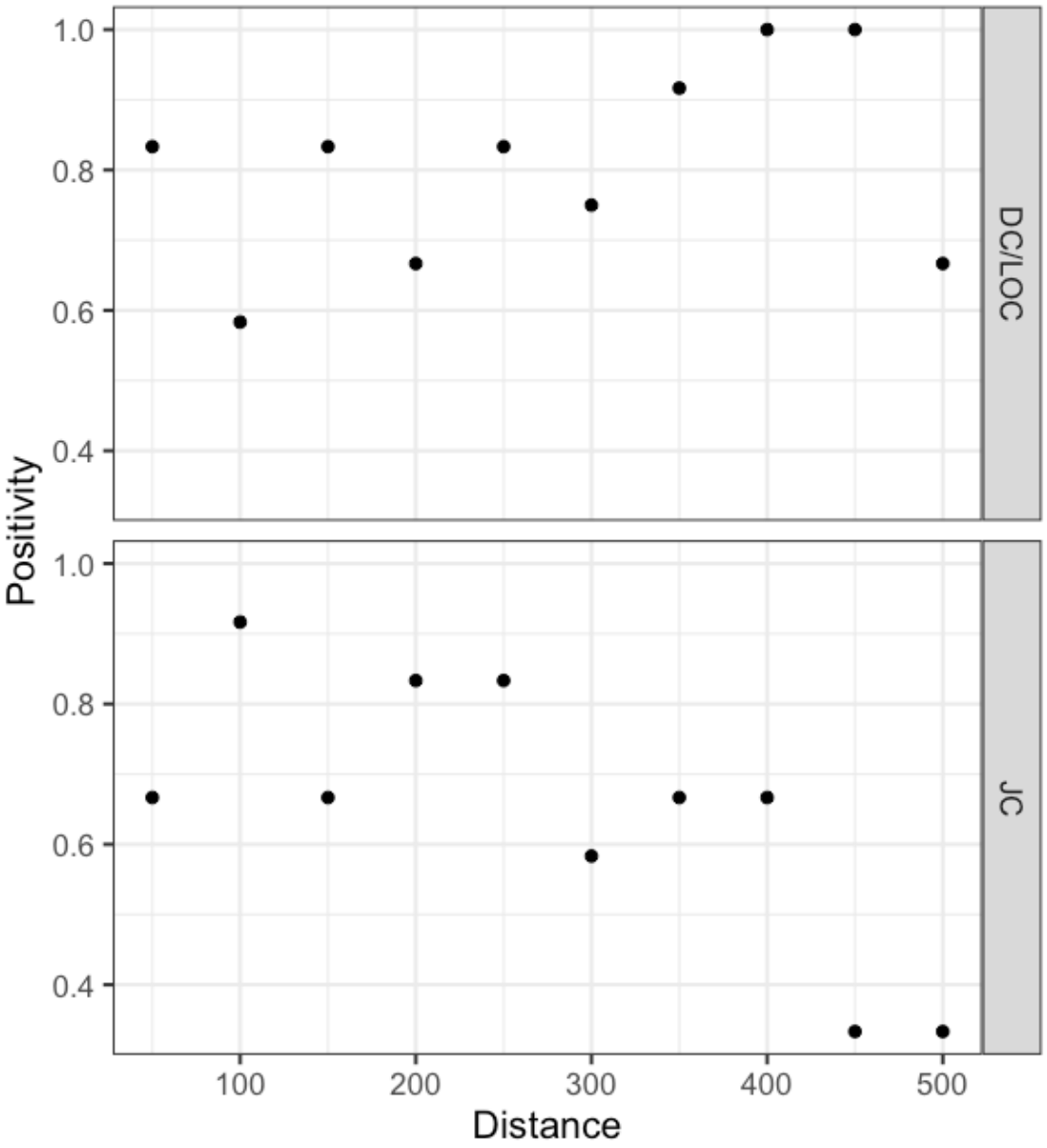

Figure A-3: Fraction of filters returning positive identification of *P. collina* DNA template as a function of distance from target population. DC/LOC and JC denote sampling sites. Exact site names are not reported given this species' status as endangered.