Guidance for Submitting Environmental DNA (eDNA) Data to the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database

Suggested Citation:

Ferrante, J.A., Neilson, M.E., Daniel, W.M., Freedman, J.A., and Hunter, M.E. 2023. Guidance for Submitting Environmental DNA (eDNA) Data to the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database: U.S. Geological Survey, Nonindigenous Aquatic Species Database, Gainesville, FL, https://nas.er.usgs.gov/eDNA/Guidance_Document.pdf. V1.2 June 2024

TABLE OF CONTENTS

INTRODUCTION

Thank you for your interest in submitting environmental DNA (eDNA) data to the U.S. Geological Survey's Nonindigenous Aquatic Species (NAS) Database. The NAS Database is an open central repository for spatially referenced accounts of introduced aquatic species. To date, it is populated by visual sightings of ~1,390 aquatic invasive species (AIS) across the United States and U.S. territories. It is our goal that the incorporation of eDNA data with traditional specimen sightings will provide more complete species distribution records to assist with national early detection and rapid response (EDRR) efforts. In this document, we discuss the justification to augment the NAS Database with eDNA data and describe the scope of the efforts and processes we used to integrate this new visual layer for eDNA data. We then outline the eDNA data approval and submission process, providing guidance for completing the necessary forms.

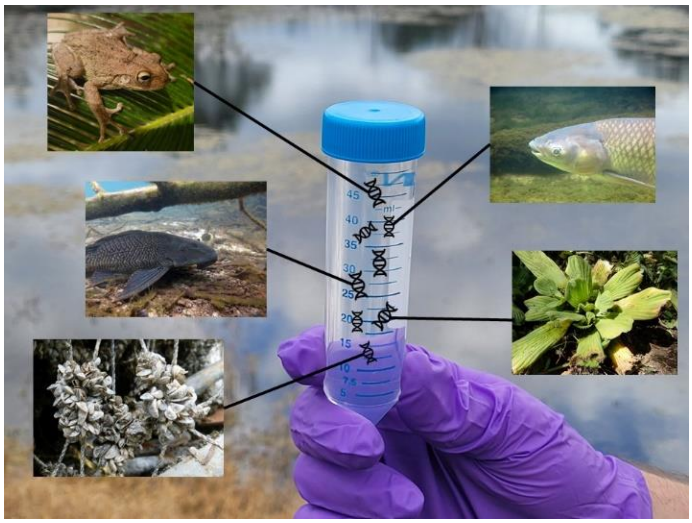


Figure 1. Environmental DNA from various species of organisms can all be captured in a single sample.

Definition of eDNA: environmental DNA (eDNA) is genetic material that has been shed by organisms into their environment. Sources of shed DNA or cells can include sloughed skin, saliva, feces, gametes, and decaying materials. Environmental DNA is often collected from water or sediment samples but can also be collected from the air.

BACKGROUND

Purpose for the incorporation of eDNA data into the NAS Database

Across the government, academia, and private industry, scientists are developing and applying novel tools for improved invasive species management and scientific decision making. The rapid development of eDNA methodologies has provided valuable tools for the biosurveillance of invasive species, increased understanding of species distribution and invasion

pathways, and estimation of relative occurrence rates (Wilcox et al. 2013, Hunter et al. 2015, Dunker et al. 2016, Hunter et al. 2019). Providing an integrated view of national eDNA and visual detections of aquatic invasive species in NAS will result in more comprehensive distribution records and could improve the response time to new incursions as part of EDRR efforts. Co-locating visual and eDNA data enables natural resource managers to make more informed resource allocation and management decisions. It also allows for the ability to harvest the data for use in secondary analyses, such as the development of risk profiles, distribution models, and the like. As data populate the database, managers will have rapid access to past and present data from multiple data sources, better informing biosurveillance and the national EDRR Framework.

Community consensus during the development of the NAS eDNA Database and data submission forms

We employed a community consensus approach to develop a process which would enable the incorporation of eDNA geospatial occurrence data to the NAS Database. We identified a set of minimum standards and best practices for the verification of eDNA data by working closely with AIS community practitioners of eDNA based detection methods and natural resource managers across government, private and academic sectors. This process was managed by a core scientific advisory panel composed of U.S. Department of the Interior invasive species eDNA scientists and members of the NAS Database. Pertinent questions regarding standards and best practices were identified from the literature and formatted into an application, which was then submitted for multiple, iterative reviews for community consensus by an expanding network of Federal, State, and university eDNA researchers, industry professionals, and invasive species managers (Ferrante et al. 2022). To submit eDNA data for inclusion in the NAS Database, contributors are first assessed via the application process and, if approved, the contributor is then invited to submit their data and associated metadata via a provided template.

Summary of the scope of the database and conditions for submission of eDNA data

NOTE: Prior to preparing data for submission, ensure that the data meet the requirements (below), and the owners of the data agree to the terms of use.

Scope: Currently, this data layer only includes eDNA detection/non-detection results from analyses targeting **nonindigenous aquatic (or semi-aquatic) species outside of their native or historic range** (not native nor terrestrial) per the NAS mission. This database is also currently limited to **probe-based quantitative PCR and digital PCR data** from **water** or **aquatic sediment** sampling (Initial Request Form, page #). This database will house the detection or non-detection results of a study, as well as the metadata for the experimental methodology, but will not function as a repository for the entire raw dataset.

- **Phase two**: We are working to incorporate **metabarcoding** data and possibly conventional/traditional/**end-point PCR** data (when paired with amplicon sequencing). In the same vein, **dye-based** (ex. SYBR) reporting chemistry will be considered in the second phase. We agree that while these additional data types are valid methods for identifying target eDNA, they are somewhat more complex to validate via our online application process, and will therefore, take more time to facilitate.

Conditions: The NAS objective is to host eDNA **data that are repeatable**. The data and accompanying metadata submitted will be displayed and available consistent with **FAIR** (Findable, Accessible, Interoperable, Reusable) guidelines.

- The verification process (Section B. Sampling and C. PCR Assay) evaluate the **experimental controls and best practices** used to produce the data. Sample blanks, controls, assay optimization and validation, LOD (limit of detection) and other quality assurance and quality controls are required, to include a minimum of 3 PCR technical replicates per reported sample.
    - To assist EDRR efforts, we have attempted to strike a balance between scientific rigor and supporting early reporting from studies that may be ongoing or not yet reported, including those with small sample sizes.
    - The stringency of the Application is intentionally highly conservative/restrictive. In the future, the application questions may be updated to maintain the rigorous vetting process while integrating the best available standards and information in the field of eDNA research. This will be a living document which can be updated to meet current methodologies and best science.

- The data presented on the NAS database viewer are submitted by various contributors and all results should be considered on a case-by-case basis. While our Application is designed to ensure contributors use best practices and employ rigorous controls in their studies, we are not able to ensure the accuracy of their data or results. Numerous variables affect eDNA presence, detection, and persistence such as species movement, seasonality, temperature, water flow, etc.; accordingly, we recommend considering the metadata associated with datasets when interpreting the results depicted in the layer.

Preparing your study: This guidance document does not provide instructions as to how to design one's study with the exception of indicating which minimum parameters are required to submit data for display on the NAS Database itself. We do note where helpful references exist and will add to these citations, as appropriate. For managers and researchers, alike, we recommend reaching out to eDNA experts during the study design process and encourage dialogue between researchers and their local AIS managers about thresholds for reporting detections in order to best inform decision-making. The Fisheries and Oceans Canada Canadian Science Advisory Secretariat (CSAS) prepared a document to facilitate communication among researchers and managers (Abbott et al. 2021). The document provides a template for reporting eDNA based data to managers in a manner that can be adapted by the researcher and manager prior to work during the experimental design phase.

THE PRE-SUBMISSION APPLICATION PROCESS

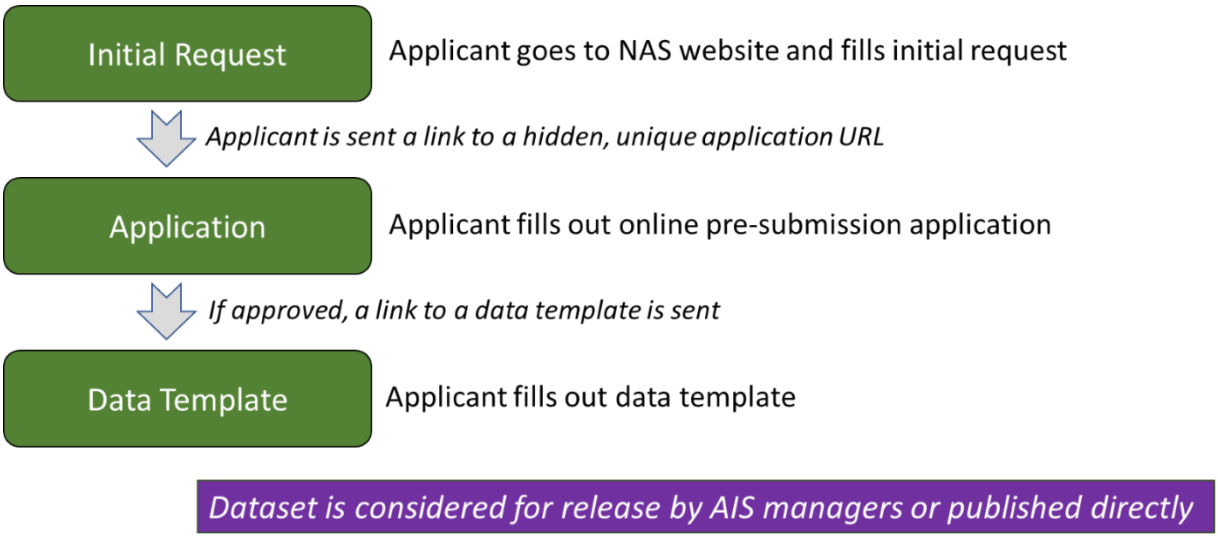


Figure 2: The pre-submission application process. This figure outlines the steps taken from an applicant's initial request to the display of eDNA data on the NAS Database.

Initial Request form

The website is structured to contain the initial online request with introductory questions that establish where the study was performed, what species were targeted, and a few other key questions to verify that the dataset is appropriate for the site (e.g., the study targets aquatic invasive species, employs hydrolysis probe qPCR or dPCR assays, etc.). This is also where the applicant indicates if their data are subject to non-preferential release rules and unable to be shared with management points of contact (POC's) prior to publication on the website. If the request meets the appropriate basic requirements listed above, the applicant is provided a link to the online application.

The full description and details are provided beginning on Page 13 of this document. *Please carefully review the data requirements for submission and display of data prior to initiating the process*.

Application

The online *pre-submission application* (hereafter Application) gives the contributor an opportunity to provide information on their experimental process, describe efforts taken to ensure proper experimental controls were employed, and detail the best practices used to avoid contamination. We advocate for the use of best practices from the literature for eDNA experimental study design, analysis, and interpretation and provide example references (Laramie et al. 2015, Goldberg et al. 2016, Hunter et al. 2017, Klymus et al. 2020). Each question is formatted as a "Yes" or "No" response with an additional space for commenting. Each question, however, must be answered in the affirmative, or the site will flag it as "not approved" and alert both the applicant and the NAS team. The NAS team can work with the applicant, if they wish, to determine how their data may become compliant with the minimum requirements. If not approved, we will inform the applicant of the reason it was denied (presumably a "No" response to a question) and work with them to see if there is a portion of their data that meets the standards, or a way to meet the standards that would allow them to re-submit. Upon approval of the Application, a *Data Submission Template* (.csv format) will be provided to be populated by the contributor along with a unique, secure URL for data submission. While the template was designed specifically for the NAS database, we sought to make the data interoperable with other databases (e.g. eDNAtlas, Darwin Core, etc.) and used existing data field terms, where possible. This includes data fields which represent the coordinates of the collected samples, taxon information, PCR results, etc. Certain fields within the template will be required, while others are optional.

The full description and details are provided beginning on Page 20 of this document. *Please carefully review the data requirements for submission and display of data prior to initiating the process*.

Data template

After applications are approved, a data template (.csv format) will be provided for the applicant to enter the data for display. The template will allow for detailed eDNA data and metadata to be uploaded to the NAS database.

The full description and details are provided beginning on Page 33 of this document. *Please carefully review the data requirements for submission and display of data prior to initiating the process*.

Communication plan

Upon receipt of the data, and in keeping with the existing stakeholder communication plan employed by NAS for visual sighting reports, a new eDNA data communication plan will be followed (Figure 2). This plan ensures that new information will be conveyed to the appropriate natural resource managers points of contact (POC) representing the geographic region. The NAS eDNA detection communication plan formally: 1) defines an updated and maintained list of natural resource management agencies and relevant job titles that will receive communications regarding an eDNA detection record; 2) describes when that information will be delivered to the POC(s); 3) determines the communication channels and methods that will be used to notify stakeholders; and 4) provides a decision tree for when the eDNA data will be made publicly available. A summary of the information will be sent to the POC, not the raw data itself. Note that this does not apply to applicants who have a non-preferential release rule, which does not allow for informing one group prior to full public data release. The eDNA communication tree is based on the current NAS database communication plan to communicate visual sightings of invasive species to pertinent aquatic invasive species (AIS) managers. Once approved for display by the POCs, or following an agreed upon embargo period, the data will be uploaded to the database layer. If the application is not approved, the applicant should be aware that details from their initial request (such as reported species detections) may still be shared with the appropriate partner agencies as to the possible presence of an aquatic invasive species.

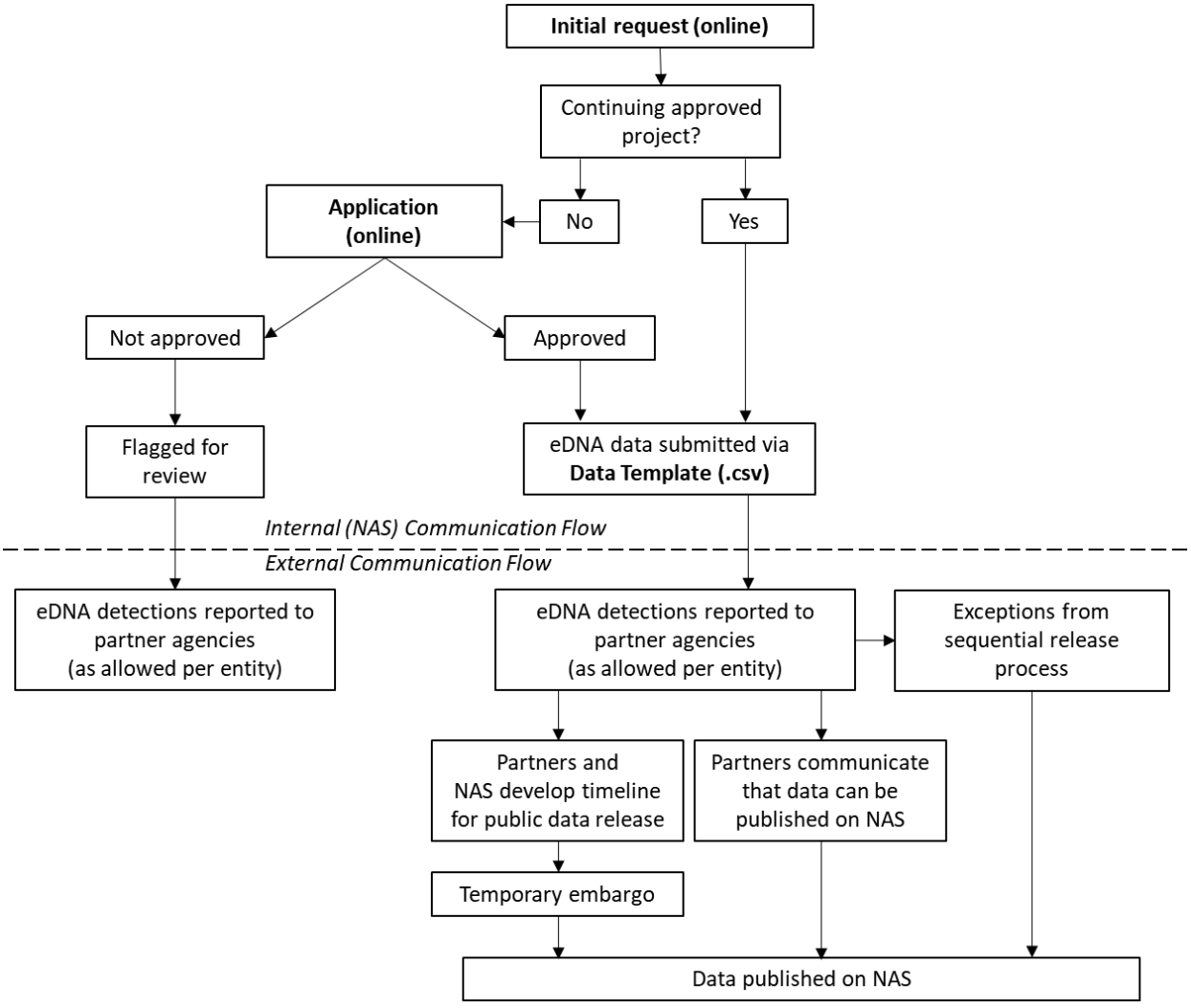


Figure 3. Communication plan from (Ferrante et al. 2022) for environmental DNA (eDNA) data submission and display on the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) database. Bold text denotes required online forms to be filled out by an applicant. Below the dotted line describes the notification of partner agencies (as allowed by the entity represented by the applicant) that a detection may have occurred in their jurisdiction. The left side of the plan shows the process for applications not approved, while the right side shows possible outcomes for those that were approved. Partner agencies include the local, Tribal, State, and/or Federal agencies with jurisdiction in the area(s) of the detection.

Display of approved data

To better inform management decisions, verified AIS eDNA data will be displayed on NAS as a separate mapping layer alongside visual sighting data. Additional information will be included for each sampling point including the eDNA methods employed to collect and produce the data.

At present, NAS displays visual sighting observations on an interactive map that allows the user to customize the base map, include various thematic overlays (e.g. watershed or Congressional district boundaries), and alternate between individual points and geographic clusters of records. Given the ability of eDNA assays to yield both positive detection and non-detection information (compared to the positive-only nature of the sighting data currently in NAS), symbology that clearly represents detections and non-detection while being visually distinct from existing visual observations is important. However, given the potential for interpretation issues around false-positive inferences defined above (i.e., detection of genetic material in a sample without visual detection of the organism at the site), it is important that the user is made aware of the nature of eDNA detection data and that such data is only presented to users after they signify their understanding. These two points lead us to use an "opt-in" model for display of eDNA detection data on the species distribution map. By default, a user will only be presented with the visual observation data when the map opens. If eDNA detection data are available for a species, a checkbox will appear in the existing species box (where users can control visibility of observation data, native range polygons, or watershed-level distribution polygons). When checked, the user will be presented with a pop-up disclaimer describing some of the limitations of eDNA detection data and how it may be interpreted, along with a distribution map (Figure 4).

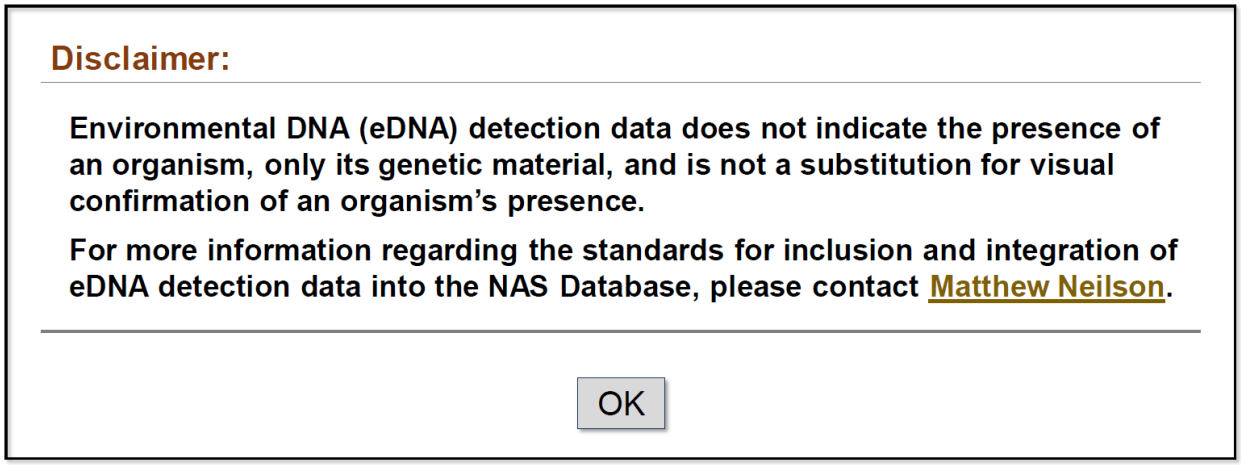


Figure 4. Image of pop-up disclaimer for eDNA data on the NAS database when that data type is selected.

The following section will guide you through the application process by showing the various forms that need to be filled out and by giving more detail about some of the questions.

INITIAL REQUEST FORM

Following the link (https://nas.er.usgs.gov/eDNA/prescreening.aspx) on the main NAS eDNA page, the first form an applicant is asked to fill out is the "Initial request form for submitting environmental DNA (eDNA) data." This Initial Request Form (IRF) is designed to ensure that the data that one wishes to submit meets the basic criteria for the current database. Following approval of the form, the database managers will provide an applicant with the Application Form appropriate for their assay type. Currently (as of Jan 2023), the database will accept data produced using only probe-based quantitative PCR or digital PCR data. As other data type entries are adapted to the database (e.g. metabarcoding data, etc.), the NAS webpage will be updated to reflect the new options.

When visiting the IRF page, the first thing you will see is a pop-up of the Paperwork Reduction Act Statement (Figure 5) as required by Federal Regulations.

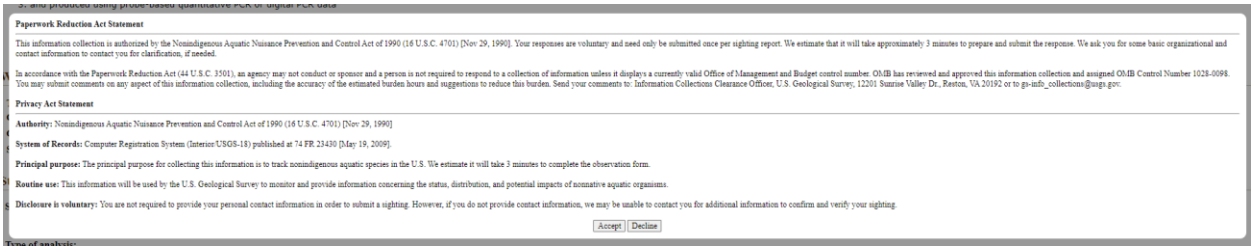


Figure 5: Image of Paperwork Reduction Act statement popup, which is displayed to users on the first visit to the Initial Request page

Text of the Paperwork Reduction Act Statement

This information collection is authorized by the Nonindigenous Aquatic Nuisance Prevention and Control Act of 1990 (16 U.S.C. 4701) [Nov 29, 1990]. Your responses are voluntary and need only be submitted once per sighting report. We estimate that it will take approximately 3 minutes to prepare and submit the response. We ask you for some basic organizational and contact information to contact you for clarification, if needed.

In accordance with the Paperwork Reduction Act (44 U.S.C. 3501), an agency may not conduct or sponsor, and a person is not required to respond to, a collection of information unless it displays a currently valid Office of Management and Budget (OMB) control number. OMB has reviewed and approved this information collection and assigned OMB Control Number 1028-XXXX. You may submit comments on any aspect of this information collection, including the accuracy of the estimated burden hours and suggestions to reduce this burden. Send your comments to: Information Collections Clearance Officer, U.S. Geological Survey, 12201 Sunrise Valley Dr., Reston, VA 20192 or to gs-info_collections@usgs.gov.

Privacy Act Statement

Authority: Nonindigenous Aquatic Nuisance Prevention and Control Act of 1990 (16 U.S.C. 4701) [Nov 29, 1990]

System of Records: Computer Registration System (Interior/USGS-18) published at 74 FR 23430 [May 19, 2009].

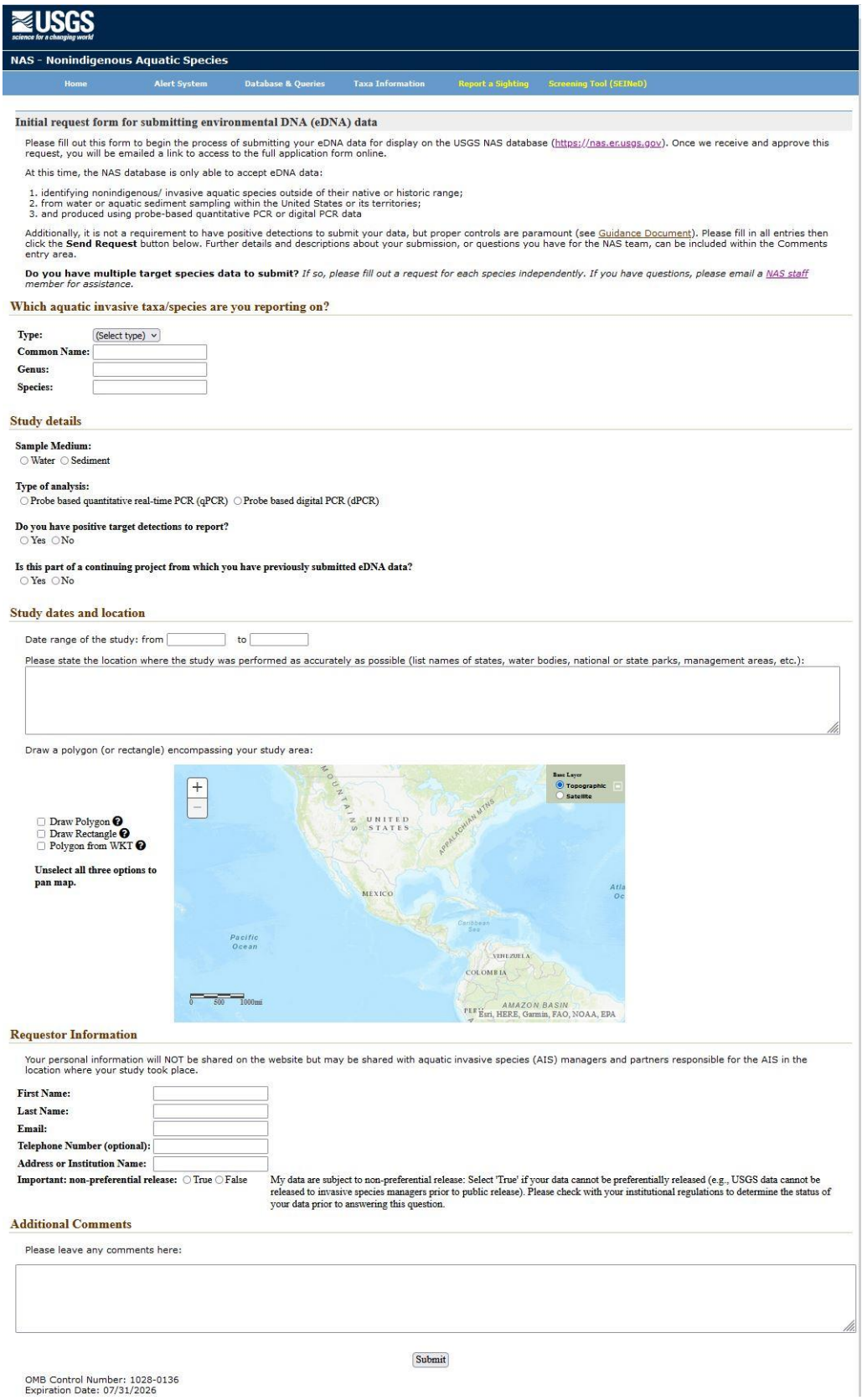
Principal purpose: The principal purpose for collecting this information is to track nonindigenous aquatic species in the U.S. We estimate it will take 3 minutes to complete the observation form.

Routine use: This information will be used by the U.S. Geological Survey to monitor and provide information concerning the status, distribution, and potential impacts of nonnative aquatic organisms.

Disclosure is voluntary: You are not required to provide your personal contact information to submit a sighting. However, if you do not provide contact information, we may be unable to contact you for additional information to confirm and verify your sighting.

By accepting the statement, the applicant will then be presented with the Initial Request form.

Figure 6: Image of the Initial Request Form



USGS
science for a changing world

NAS – Nonindigenous Aquatic Species

Home | Alert System | Database & Queries | Taxa Information | Report a Sighting | Screening Tool (SEINeD)

Initial request form for submitting environmental DNA (eDNA) data

Please fill out this form to begin the process of submitting your eDNA data for display on the USGS NAS database (https://nas.er.usgs.gov). Once we receive and approve this request, you will be emailed a link to access to the full application form online.

At this time, the NAS database is only able to accept eDNA data:

1. identifying nonindigenous/ invasive aquatic species outside of their native or historic range;
2. from water or aquatic sediment sampling within the United States or its territories;
3. and produced using probe-based quantitative PCR or digital PCR data

Additionally, it is not a requirement to have positive detections to submit your data, but proper controls are paramount (see Guidance Document). Please fill in all entries then click the **Send Request** button below. Further details and descriptions about your submission, or questions you have for the NAS team, can be included within the Comments entry area.

Do you have multiple target species data to submit? *If so, please fill out a request for each species independently. If you have questions, please email a NAS staff member for assistance.*

Which aquatic invasive taxa/species are you reporting on?

Type: (Select type)
Common Name:
Genus:
Species:

Study details

Sample Medium:
○ Water ○ Sediment

Type of analysis:
○ Probe based quantitative real-time PCR (qPCR) ○ Probe based digital PCR (dPCR)

Do you have positive target detections to report?
○ Yes ○ No

Is this part of a continuing project from which you have previously submitted eDNA data?
○ Yes ○ No

Study dates and location

Date range of the study: from to

Please state the location where the study was performed as accurately as possible (list names of states, water bodies, national or state parks, management areas, etc.):

Draw a polygon (or rectangle) encompassing your study area:

☐ Draw Polygon
☐ Draw Rectangle
☐ Polygon from WKT

Unselect all three options to pan map.

Requestor Information

Your personal information will NOT be shared on the website but may be shared with aquatic invasive species (AIS) managers and partners responsible for the AIS in the location where your study took place.

First Name:
Last Name:
Email:
Telephone Number (optional):
Address or Institution Name:
Important: non-preferential release: ○ True ○ False

My data are subject to non-preferential release: Select 'True' if your data cannot be preferentially released (e.g., USGS data cannot be released to invasive species managers prior to public release). Please check with your institutional regulations to determine the status of your data prior to answering this question.

Additional Comments

Please leave any comments here:

Submit

OMB Control Number: 1028-0136
Expiration Date: 07/31/2026

The Initial Request Form is then displayed as a web form. The collated text and questions are copied below (boxes) with some additional descriptions and definition narratives (*italicized* within the boxes) to aid in the completion of the form.

The first section of the IRF describes the currently allowable data for the database:

Initial request form for submitting environmental DNA (eDNA) data

Please fill out this form to begin the process of submitting your eDNA data for display on the USGS NAS database (https://nas.er.usgs.gov). Once we receive and approve this request, you will be emailed a link to access the full application form online.

At this time, the NAS database is only able to accept eDNA data:

1. identifying nonindigenous/invasive aquatic species outside of their native or historic range;
2. from water or aquatic sediment sampling within the United States or its territories;
3. and produced using probe-based quantitative PCR or digital PCR data

Additionally, it is not a requirement to have positive detections to submit your data, but proper controls are paramount (see Guidance Document). Please fill in all entries then click the **Send Request** button below. Further details and descriptions about your submission, or questions you have for the NAS team, can be included within the "Comments" entry area.

Do you have multiple target species data to submit? If so, please fill out a request for each species independently. If you have questions, please email a NAS staff member for assistance.

The second section allows for reporting of the target species:

Which aquatic invasive taxa/species are you reporting on?

Type: (Select type - *dropdown*) *from Amphibian, Crustacean, Fish, Marine Fish, Mollusk, Plant, Reptile, Other*

Common Name: *list all known; text box provided*

Genus: *text box provided*

Species: *if known; text box provided*

The third section allows for reporting of details that relate to the study. This section will adapt as new data types are allowable for submission.

Study details

Sample Medium: *Choose water or sediment*

Type of analysis: *Choose Probe based quantitative real-time PCR (qPCR), or Probe based digital PCR (dPCR)*

Do you have positive target detections to report? *Choose Yes or No*

Is this part of a continuing project from which you have previously submitted eDNA data? *Choose Yes or No*

The fourth section allows for reporting geospatial and temporal information related to the dataset.

Study dates and location

Date range of the study: from _____ to _____; *text box provided*

Please state the location where the study was performed as accurately as possible (list names of states, water bodies, national or state parks, management areas, etc.): *a text box is provided for your answer.*

Draw a polygon (or rectangle) encompassing your study area: *Choose the type of method you will use to outline your study area - Draw Polygon, Draw Rectangle, Polygon from WKT. Unselect all three options to pan map.*

The final sections allow for the provision of contact information, selection of non-preferential release status, and any additional comments by the applicant. Contact information is used by the

NAS managers to communicate with the applicant but will not be shared via the database directly.

Requestor information

Your personal information will NOT be shared on the website but may be shared with aquatic invasive species (AIS) managers and partners responsible for the AIS in the location where your study took place.

First Name: *A text box is provided*

Last Name: *A text box is provided*

Email: *A text box is provided*

Telephone number (optional): *A text box is provided*

Address or Institution name: *A text box is provided*

Important: Non-preferential release

My data are subject to non-preferential release: *Select True if your data cannot be preferentially released (i.e. USGS data cannot be released to invasive species managers prior to public release). Please check with your institutional regulations to determine the status of your data prior to answering this question.*

Additional comments

Please leave any comments here: *A text box is provided*

Following submission of the initial request, the form is reviewed for the appropriate basic requirements (e.g. your study targets aquatic invasive species, employs hydrolysis probe qPCR or dPCR, etc.). If approved, the applicant is provided a link to the online Application. Details for filling out the Application are in the next section of this Guidance Document.

APPLICATION FORM

The following YES or NO questions emphasize the use of best practices and proper experimental controls to avoid contamination and to ensure accurate reporting of results. Due to the rapidly changing science, and public dissemination of results, stringent requirements were adopted for these data standards. In addition to the standards indicated in the questions below, we encourage all researchers to apply general best practices for avoiding contamination in their study designs, such as the use of gloves for handling samples and single use consumables, as well as dedicated sample collection gear and lab equipment that is sterilized appropriately between samples. This application will adapt and evolve as standards and best practices are further refined.

Note that each of the following application questions require a "YES" response for a submission to move forward. An applicant is asked to provide the answers to the questions strictly related to the specific data they wish to submit to the database. Failure to meet the data standards (*i.e.*, answering "NO" to any of the questions) will result in your data submission failing the quality check. For applicants who find they cannot answer a question in the affirmative, we welcome them to contact the NAS staff directly (https://nas.er.usgs.gov/about/staff.aspx) to see if they can be of assistance or provide guidance in meeting the requirement.

If a fee-for-service lab processed the eDNA samples, please enlist their help to complete the application. We currently do not accept fully automated sample collection (e.g., autonomous filtering and PCR instruments). **By submitting data, an applicant agrees to follow the general scientific ethical guidelines of consent for submission from coauthors and/or partners and funders and reporting all data and information truthfully and without fabrication.**

Following are the Application questions with additional descriptions and definitions **provided in boxes**. Terms that are underlined can be found in the glossary.

A. Basic Study Information

A.1 Was either probe-based quantitative PCR (qPCR) or digital PCR (dPCR) performed for all sample data being reported? Please indicate which: ☐qPCR or ☐dPCR (radio-select)

Description: Only data produced via hydrolysis, probe-based (e.g. TaqMan, etc.) quantitative PCR or digital PCR assays are applicable. Non-applicable data include data produced using methods such as non-specific, fluorescent dye-based qPCR (e.g., SYBR Green, EtBr, LAMP, etc.), metabarcoding, or conventional/end point PCR.

A.2 Can the collection date be provided for each water or sediment sample for which data are being reported?

Description: The date that the sample was collected, preferably using the ISO 8601 date format of Full Year-Month-Day (e.g., 2019-07-26), must be available for each sample reported.

A.3 Can geographic coordinates be provided for each sample collection station?

Description: A station is defined as a geospatially distinct sampling location where a water or sediment sample is obtained (see Box 1), as opposed to a "site" which is more broadly defined (e.g., river, canal, hill, etc.). Often multiple or replicate stations are sampled within a site. For each station, you must be able to provide the latitude and longitude (WGS84 datum) for each water or sediment sample to a minimum of three decimals (e.g., 29.725, -82.418). For samples obtained using a moving filtration system, such as a backpack, the starting coordinate is requested.

BOX 1: Defining geospatial locations

Station: A geospatially distinct location where a sample is obtained. The most precise location generally labeled for sampling. A station will have unique coordinates relative to other stations. Many stations can be found within a site (i.e. spatial replicates), and multiple samples may be obtained at a single station. When sampling along a transect, each location that a sample is collected is a separate station.

Site: A geographic feature that is independent of another site, such as a different habitats or minor ecosystems (e.g., lakes, ponds, rivers, creeks, individual canals).

Region: Larger geographic units that encompass numerous sites – a park or preserve, drainage basin, county, etc.

B. Sampling and Processing

Sampling

B.1 Can the volume of the water sample or mass of sediment sampled be provided?

Description: A water volume (in liters) or sediment mass (in grams) must be reported for each sample leading to an individual data point.

B.2 Were the initial field samples protected from light and (1) immediately preserved using a DNA stabilizing buffer (i.e., EtOH for final 80% concentration) and then stored on ice/refrigeration (4°C) until processed; or (2) frozen (0°C minimum) within 24 hours of collection? Alternatively, were the samples (3) filtered on site and then preserved (using stabilizing buffer, desiccant, *etc*.) for analysis off-site, or (4) filtered and immediately analyzed on-site? Answer Yes if ANY of the methods were employed.

Please indicate which method(s) was/were used: (radio-select – can have multiple answers)

☐ stabilized with preservative on site and stored at 4°C

☐ frozen for future processing or filtered within 24 hours

☐ filtered and/or preserved (e.g., desiccant, freezing, etc.) on site

☐ PCR analyzed on site

Description: Evidence shows rapid degradation of eDNA from water or sediment samples at warmer temperatures (Strickler et al. 2015). A maximum short-term storage temperature of 4°C is needed to preserve eDNA yield and quality, prior to and after concentration (i.e., filtration or centrifugation). If no preservation method listed was used, samples must have been analyzed through the PCR step on location for approval of the application.

Contamination Controls

Negative controls, or "blanks," are required for approval of the application. The next questions concentrate on controls which account for possible contamination sources. For each type of control, you will need to verify that the PCR did not amplify DNA. This result applies to negative controls collected and associated with the specific samples you wish to report.

B.3 Field equipment blanks: Was at least one equipment blank (DNA-negative water or sediment) prepared at each site (per day) using the field sampling methods to ensure neither the equipment, nor site, were contaminated by DNA introduced prior to the field effort? Please indicate the frequency and timing of field blank preparation (*i.e.*, daily, upon arrival at each new site, after finishing at each station, for each sampling event at a station, *etc.*) in the comments.

*Description: We **require**, at a minimum, one field blank at each site (site broadly defined as a river, canal, hill, etc.), and if multiple days at a site, a field blank must be collected for each day. We recommend doing so prior to the beginning of sampling. Please indicate the frequency (daily, at each site, at each station, etc.) of field blank collection in the **comments**. A field blank is collected to ensure the site or equipment is not contaminated by DNA introduced from the lab, or a previous sampling site, etc. It involves pouring molecular grade water or dispensing DNA-free sediment into a collection receptacle (e.g., bottle, tube, bag, etc.) in the field. When using a field filtering apparatus (mentioned above), this would mean pulling DNA-free water through the equipment system and collecting the blank as if it were a sample. Optimally, this is done before field sampling begins and using decontaminated equipment.*

B.3a Did all the field equipment blanks result in no amplification of DNA?

Description: Please verify that the PCR did not amplify target or non-target DNA. This result applies to blanks collected and associated with the specific samples you wish to report.

B.4 blanks: Were DNA-negative water or sediment blanks processed using the same eDNA concentration (isolation from substrate) methodology (*i.e.,* did you filter/centrifuge negative controls) every 24 samples.

Description: A method blank is collected to ensure the equipment used for isolation is not contaminated as a result of processing. It involves preparing and analyzing DNA-free water or sediment using the same protocol as the field samples including the concentration equipment (i.e., filtration systems, centrifuge, etc.). Concentration, or isolation, here refers to the step wherein the biological material is concentrated or separated from the medium (water or sediment) and does not refer to the extraction of the DNA from the biological material (see Q. B.5).

B.4a Did all the concentration-blank(s) result in no amplification of DNA?

Description: Please verify that the PCR did not amplify target or non-target eDNA. This result applies to blanks collected and associated with the specific samples you wish to report.

B.5 DNA extraction blanks: Were DNA-negative water or sediment samples processed using the same eDNA extraction (purification) method and/or kits to test for contamination every 24 samples?

Description: Here, potential contamination of the eDNA extraction step is tested by running a target DNA-negative water (e.g., distilled) or sediment sample, using the DNA extraction kit/method.

B.5a Did all the extraction blank(s) result in no amplification of DNA?

Description: Please verify that the PCR did not amplify target or non-target eDNA. This result applies to blanks collected and associated with the specific samples you wish to report.

Processing Methods

B.6 Was a peer-reviewed, published method for extraction (purification) of eDNA samples used in the study? Please select the product type (radio-select) and include a reference and/or commercial product name in the comments.

☐ Silica column

☐ Phenol-chloroform

☐ Enzyme

☐ Bead

☐ Other published method (MUST include reference in comments)

Description: At this time, only extraction methods validated for eDNA samples published in peer-reviewed journal articles are accepted. Studies which use novel, unpublished methods are unable to be accepted by our process at this time. Please include a peer-reviewed, published reference in the comments for the method you employed, and/or the specific product name.

C. PCR Assay

Validation and Optimization

PCR assay validation

In this section, validation refers to the detection or amplification of target eDNA and non-detection or amplification of non-target eDNA. For some questions, you will be asked to test in-house efforts versus reporting the results of a previously published study used as the reference for your study. We require that assay validation based on a laboratory different than the submitting lab are applicable ONLY if they tested species (*in silico*) or samples (*in vitro/situ*) from your study area at the regional scale (defined as a larger geographic unit encompassing numerous sites, such as a park or preserve, drainage basin, county/state, *etc.*). Otherwise, we require that these validation steps be assessed by the lab analyzing the experimental samples.

Please cite the referenced study for the assay in the comments. At this time, only qPCR or dPCR methods validated for eDNA samples and published in peer-reviewed journal articles are accepted. Studies which use novel, unpublished methods are unable to be accepted at this time.

However, if the novel protocol gets published, then it may be admissible in the future.

C.1 Has the specificity of the PCR primers and probes been validated to exclude non-targeted taxa *in silico*?
If yes, are you reporting based on (radio-select):
☐published information from another lab
☐results from lab that performed the assay producing the data for submission

Description: Were in silico tools/software used to validate the specificity of primers and probes to the taxa's target sequence? Any appreciable homology in non-targeted species that may be in the study area (at a regional scale) should be noted in the comments.

C.2 Has the PCR assay been validated for the exclusion of non-targeted taxa specificity *in vitro*?
If yes, are you reporting based on (radio-select):
☐published information from another lab
☐results from lab that performed the assay producing the data for submission

Description: Was this assay tested against genetic samples (e.g., DNA extracts, etc.) from non- target taxa? This question seeks to ensure the assay did not cross-react with an organism that could occur within the system being sampled (regional scale). This is often accomplished by performing a PCR (endpoint, qPCR or dPCR) to test genetic samples from closely related, non- target species of the same genera and non-related genera. If this test resulted in a positive detection for a non-target species, please list the non-target species (scientific names) in the comments.

C.3 Has the PCR assay been validated for taxa specificity *in situ* against co-occurring, non-target taxa?
If yes, are you reporting based on (radio-select):
☐published information from another lab
☐results from lab that performed the assay producing the data for submission

Description: Please indicate whether water or sediment samples from a site within the study region where the target taxa are known to be absent have been tested using the assay and that the confirmed target taxon was not detected. For this control, the sample is processed like a blank and assayed with an internal positive control (IPC) not derived from the target species to ensure the PCR reaction occurred, but that no non-specific amplification occurred.

C.4 Has the PCR assay been validated for target taxa *in situ* using target eDNA positive water or

sediment samples obtained from the geographic region of your study?
If yes, are you reporting based on (radio-select):

☐published information from another lab

☐results from lab that performed the assay producing the data for submission

Description: This question addresses the possibility of false negatives based on genetic differences within a species. For example, using an assay designed for target taxa in the Great Lakes to test for the same target taxa in the Pacific Northwest would not meet this criterion. Here we ask if the qPCR or dPCR assay was tested by sampling in an area where the target taxa are known to be present. Alternatively, this can be done by spiking water or sediment from the study area with target taxa DNA and observing a positive detection for the target DNA. Spiking the sample is acceptable for samples from study areas where a species has not yet invaded or in rare cases, suitable samples of similar habitat type are not accessible.

C.5 Has the PCR assay been validated *in situ* to test for inhibitors and other environmental cofactors in the study area that may affect PCR amplification of your target eDNA?
If yes, are you reporting based on (radio-select):

☐published information from another lab

☐results from lab that performed the assay producing the data for submission

Description: Although the chemical and physical makeup of a water body or area of sediment can change daily, seasonally, etc., this generally validates the assay functionality in the area being sampled. One way to assess inhibitors is to run tests of water or sediment samples with internal positive controls (IPCs) to determine whether inhibitors may be affecting amplification.

PCR assay optimization

The assays used to produce the data must meet specific standards from the Minimum Information for Publication of Quantitative Real-Time Experiments (MIQE) or digital PCR experiments (dMIQE) to include standard curves and serial dilutions within the acceptable MIQE parameters (Bustin et al. 2009; Huggett et al. 2013). The NAS objective is to host eDNA data that are repeatable. Standard curve data provide critical information regarding repeatability and quality assurance. For lab-based qPCR analysis, a standard curve must be run at least once with the same protocol and equipment (qPCR machine, reagents, *etc*.) used to analyze the samples. If a commercial field qPCR system (instrument, reagents, *etc*.) was used, the company's validation metrics for that assay can be used to answer the questions that follow.

While the following metrics (C.6, C.6a-C.6b) do not directly apply to dPCR assays, we still require that you report the metrics based on a published work that validated the primers/probes

for the dPCR assay you used (please cite your reference(s) in the comments). If such a publication does not exist, at this time we require the applicant perform a validation qPCR standard curve in their own lab and use the results to answer the questions below.

C.6 Were standard curves (qPCR) or dilution series (dPCR) run using a minimum of three technical replicates per standard/dilution with a minimum of five standards/dilutions covering the expected dynamic range, OR can you report such metrics from an approved source (see description)?

Description: A minimum of three (amplified) replicates for each of five standards are required for statistical analysis of the curve following MIQE guidelines (Bustin et al. 2009). If data are produced via a qPCR assay, a standard curve must be run with the same system and reagents assessing the experimental samples at least once prior to analyzing samples for reporting below. It is not required that standard curves be run with each set of samples (i.e., on each plate, etc.). If multiple curves are run, the minimum requirements below would apply to the dataset associated with each curve. An appropriate material used for standards would be custom or synthetic DNA, to include gene fragments, long oligonucleotides, plasmids, etc.

C.6a Was the qPCR assay efficiency between 90 and 110%?
For dPCR, are you reporting based on (radio-select):

☐published information from another lab (please cite reference in comments)

☐results from your own lab

☐I did not run dPCR

Description: Efficiency values are calculated from the standard curve and should fall within the 90-110% range. Efficiency is a measure of how effectively the assay amplifies the target sequence. Optimally this value is as close to 2-fold each cycle (C_q), or 100%. You are welcome to enter your efficiency value in the comments.

C.6b Was the R^2 of the qPCR standard curve ≥ 0.95?
For dPCR, are you reporting based on (radio-select):

☐published information from another lab (please cite reference in comments)

☐results from your own lab

☐I did not run dPCR

Description: R^2 is the coefficient of determination (i.e., goodness-of-fit) of a linear regression model. This is calculated from the standard curve, generally by the qPCR software or by a statistical analysis program. The closer the value is to 1.00, the better the regression is predicting the approximate real data points.

C.6c Was the qPCR assay's limit of detection (LOD) determined?
For dPCR, are you reporting based on (radio-select):

☐published information from another lab (please cite reference in comments)

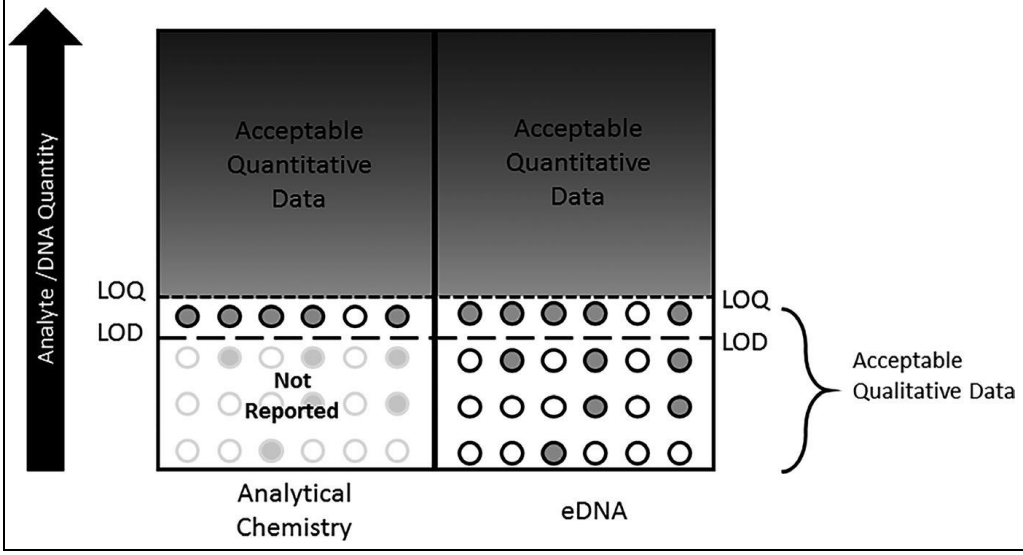
☐results from your own lab

☐I did not run dPCR

Description: The limit of detection is the lowest standard concentration with at least a 95% detection rate across all replicates (see Guidance Document). For lower concentration standards (those that fall within the double digits), it is recommended to run 10 or more replicates to ensure accuracy, as more replicates may help to decrease the effective LOD (Klymus et al. 2020).

BOX 2: Limit of Detection and Limit of Quantitation

For qPCR studies, Klymus et al. 2019, recommended that a limit of detection and a limit of quantitation should be calculated prior to interpreting the results of unknown samples. Due to the sensitivity of qPCR methods, amplification of DNA at low concentrations can occur. However, to ensure reproducibility, the LOD should be the lowest concentration with at least a 95% detection rate and the LOQ should be the lowest concentration with a coefficient of variation below 35%.



C.7 Were negative controls, or no-template controls (NTCs) run with each batch of samples or PCR plate?

*Description: It is required that negative controls (i.e., NTCs, instrument blanks) are analyzed with **each** batch of samples (i.e., each plate, run, etc.). This is appropriately accomplished if DNA-free water or buffer is used to replace template (samples, standards, etc.) in a **minimum** of two reactions (for validation by reproducibility of results) during the analysis of each batch of samples. For the purposes of this application, a batch could mean one 96-well plate for each batch of samples (samples collected from the same site on the same day) to account for reagent differences or potential on-site contamination. For assays not performed on a plate, such as field qPCR systems, a minimum of two negative controls may be analyzed across two separate runs (as some systems only analyze a few reactions per run).*

C.7a Did the negative controls result in no amplification of DNA?

*Description: Please verify that the negative control PCR reaction did not amplify target or non- target eDNA. This result applies to negative controls analyzed with each specific batch of samples you wish to report. **Note:** To allow for acceptance of the data within a batch, all negative controls associated with the batch must result in no amplification. At this time, alternative ways of correcting for amplification in negative controls is not allowed. Unallowed protocols include correction factors such as subtracting positive detection copy numbers from the experimental samples (often defined as 'limit of blank').*

C.8 Were two or more positive controls run for each PCR assay (each plate or run) on experimental samples? Please identify which type. (radio-select):

☐Genomic DNA

☐PCR amplicon

☐Custom/ Synthetic DNA (i.e., gene fragment, gBlocks, oligonucleotide, plasmid, etc.)

*Description: It is required that a **minimum of two** positive controls are run with **each** batch of samples (i.e., each plate, PCR, etc.). This is appropriately accomplished if a standard curve is included in the run, or by the addition of target DNA (e.g. from a tissue sample) to several wells (**minimum of two**) of the plate/strip/etc. to verify PCR amplification occurred. Genomic DNA could be extracted from the tissue, blood, etc. of the target organism. PCR amplicons and synthetic DNA strands should have at least five nucleotides 5' of ('flanking' or 'upstream of') the binding sites of the forward and reverse primers (e.g., the DNA fragment should not begin or end at the first nucleotide of the primer binding site). For assays not performed on a plate, such as field qPCR systems, the minimum two positive controls must be run with each batch of samples (samples collected from the same site on the same day) to account for reagent differences or local sample source variation (pH, inhibitors, etc.), but do not all necessarily need to be analyzed in the same PCR.*
***For dPCR**: If dPCR was run, the positive control should have been run to function as a*

calibrator (i.e., a DNA control containing the target sequence) (Huggett et al. 2013). This positive control should be used at the same volume and concentration throughout the study and must have been run on each plate, or with each round of analyses, as appropriate. Each reaction should yield a consistent number of copies for the calibrator/ positive control.

C.8a Did all positive control samples associated with the data you wish to submit result in (qPCR) amplification or did your positive control (calibrator) yield similar copy numbers and amplitude from dPCR run to run?

Description: Please verify that the qPCR/dPCR did amplify target DNA. This result applies to positive controls analyzed with each specific batch of samples you wish to report. If running dPCR, please verify that similar copy numbers and amplitude were observed from run to run.

C.9 Were at least three PCR technical replicates run for each sample in the PCR analysis?

*Description: A **minimum of three technical replicates per sample** (at this time) are required to ensure repeatability of results, reduce the chances of false negatives, and to accurately report true positives.*

D. Reporting

D.1 Are the data reportable for each sample in one of these standardized formats: Qualitative detection (non-detection or detection) or quantitative detection (concentration in copies/Liter for water or copies/gram for sediment)? Please indicate which or both (radio- select – can have multiple answers).

☐Qualitative

☐Quantitative (you must report the limit of quantification in the comments)

Description:
***Qualitative detection**: All reported detections must be replicated in a minimum of 2 technical replicates in each of two or more station replicates. This means that a minimum of two samples need to be taken at each station, and two technical replicates from each sample are needed to indicate a detection (e.g detection cannot be made with only one reaction/replicate). Additionally, for qPCR, detection must occur with value(s) above the Limit of Detection (LOD).*
***Quantitative detection:** Must first meet the qualifications for quantitative detection. For qPCR, the concentration value of at least two (**2) technical replicates** with positive detection must be above the assay limit of quantification (LOQ). This is defined as the lowest standard concentration that resulted in less than 35% coefficient of variation (Klymus et al. 2020). For dPCR at least two (2) technical replicates must have one or more positive droplets with an*

amplitude within the range of the positive control/ calibrator run on that plate. Calculations should use the metrics: copies/Liter for water or copies/gram for sediment.

***Non-detection**: For qualitative or quantitative reporting, this is a result of no detections of target eDNA among all PCR replicates.*

***Note about Inconclusive detection (see** Guidance document**): Defined as a detection in only 1 of the PCR technical replicates for a station. These data will be permissible to submit with your dataset but will not be displayed on the map viewer. They will be retrievable in the table format with the rest of your submitted data.*

Box 3. Determining the type of eDNA detection from your results

The eDNA of a target species may be detected along a broad range of concentrations depending on variables such as the target species abundance, sample volume, method used to extract the eDNA from the medium, and/or levels of environmental inhibitors.

Detection – do we report like a diagnostic assay?

Context matters: What happens when eDNA is present in a sample, but the organism is not present in the area the sample was taken? In the context of detecting an organism itself, a positive eDNA detection could be considered a false positive detection for the presence of the species. However, in the context of eDNA detection, it is a positive detection for the target DNA, since the genetic material was in fact present. In the scenario where we don't know if the organism was present, the eDNA detection may suggest the organism either was present, or the eDNA has moved into the sampling area from an area where the organism resides. Here is where the controls, study design and repeated sampling become paramount;field negatives, lab negatives and positives, and replicates help to ensure accurate reporting of detections. Doing so helps address manager concerns, ensures accurate interpretation, and avoids negative public perception and inappropriate allocation of limited resources. For quantitative detection reporting, dPCR target droplets in each reaction must have an amplitude within the positive control droplet amplitude range run for that plate, and for qPCR, the threshold (Cq) value of each putative detection must be above the assay LOQ. Concentration must be reportable based on the effective quantity tested (T. M. Wilcox et al. 2018) as copies/Liter or copies/gram (adjust equation).

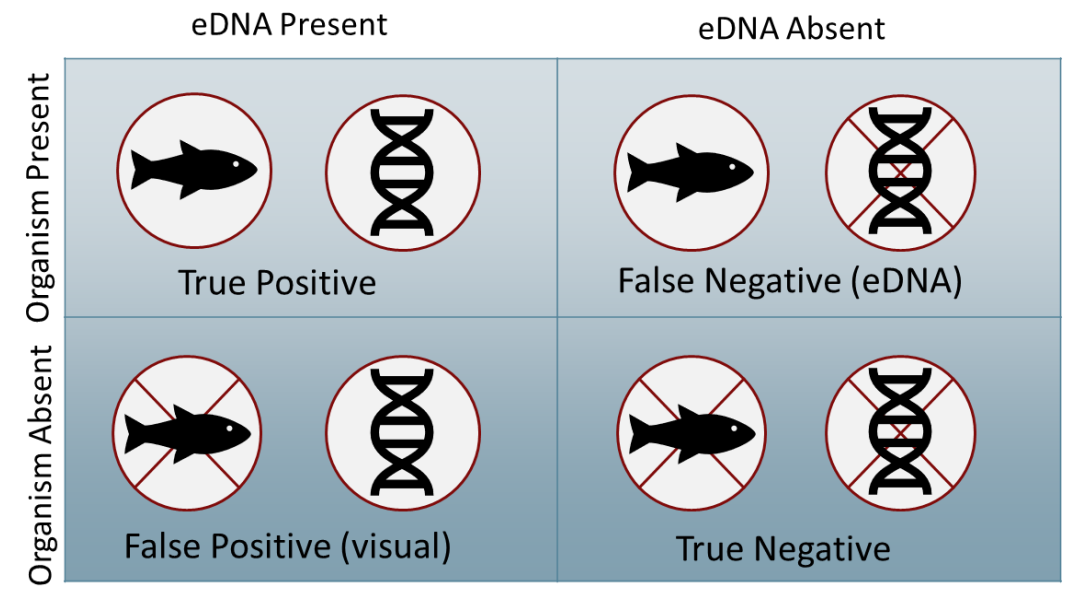

eDNA Present
eDNA Absent
Organism Present
Organism Absent
True Positive
False Negative (eDNA)
False Positive (visual)
True Negative

Figure 7

DATA TEMPLATE

Template

												Technical replicate average				
Persistent ID	Scientific name	Common name	Station ID	Replicate	Sampling Date	Latitude	Longitude	Sample type	Sampling unit	Sample amount	Assay type	PCR technical Replicates	Avg Cq	Positive detects/well	Accepted droplets/well or partition	Detection Status	Copies /L or /g
Your_PI-0001	Genus species	Common name	Station 1	A	2021-01-18	25.3819	-80.6094	Water	Liters	1	qPCR	5	31.6	N/A	N/A	Detect	1004
...	Genus species	Common name	Station 2	B	2021-01-18	25.3819	-80.6094	or	or	1	or	5	N/A	0.1	20043	Inconclusive	N/A
Your_PI-0003	Genus species	Common name	Station 3	C	2021-01-18	25.3819	-80.6094	Sediment	grams	1	dPCR	5	N/A	0	20043	Non-detect	N/A
			^ Controls (not displayed on map)									(min 3)	for qPCR		for dPCR		above LOQ
			~ Positive control/ dPCR calibration control														
			~ Negative control (NTC)														
			~ Field equipment blanks														
			~ Method blanks														
			~ Extraction blanks														

Figure 8: Data template image

The data template is a .csv file which is used to provide key information from the applicant to the database. Below are the headers with a description of the information being requested.

Column Header	Description of requested information
Persistent ID	Unique identifier (ID) for data point from the applicant
Scientific name	Provide the most currently agreed upon scientific name (genus and species) of the target organism.
Common name	Provide the common name(s) for the targeted species (if applicable)
Sample name	(e.g., Station ID or one of the controls) Provide any label used to identify the station or provide labels for various controls used in the PCR (not to be displayed on the map) to include: qPCR positive control/ dPCR calibration control (separate from standards, or provide selected standards used as positive control), Negative control (NTC), Field equipment blank, Method blank, Extraction blank.
Replicate	Provide a replicate identifier if replicates were taken at a station (alpha-numeric per your discretion)
Sample category	Indicate here whether this row contains an eDNA sample or one of the various controls used in the PCR. Possible values are ‘Sample’, ‘Positive Control’, ‘Negative Control – NTC’, ‘Field Equipment Blank’, ‘Method Blank’ and ‘Extraction Blank’.
Sampling Date	Enter the date that the sample was collected in *ISO 8601* 4-digit year – two-digit month – two-digit day (####-##-##) format.
Latitude	Provide the latitude (WGS84 datum) of the station to a minimum of three decimals (e.g., 29.726, ….).
Longitude	Provide the longitude (WGS84 datum) of the station to a minimum of three decimals (e.g., …., -82.419).
Sample type	Select if sediment or water was sampled
Sampling unit	Select if liters or grams were used as the unit of measure for the sample amount.
Sample amount	Enter the measured amount of sample. Note – just enter the value here as the unit is identified in the previous column.

Assay type	For quantitative, real-time PCR, select qPCR and for digital PCR, select dPCR.
PCR technical Replicates: (min 3)	Enter the number of PCR technical replicates analyzed to produce the result. A minimum of 3 per sample is required for publishing the data on the database.
Avg Cq (for qPCR)	Enter the average quantification cycle (Cq; sometimes Ct – threshold cycle or Cp – crossing point) of the technical replicates (for qPCR only).
Avg Positive detects (for dPCR)	Enter the average number of positive detections between the wells of the technical replicates (for dPCR only).
Avg accepted droplets (for dPCR)	Enter the average number of accepted droplets between the wells of the technical replicates (for dPCR only).
Detection Status	Select the detection status for this sample (replicate). o "Non-detection" meaning no detections of target eDNA among all PCR technical replicates. o "Detection" meaning either qualitative or quantitative detection was made among two or more station or PCR replicates (not just in one reaction/replicate), and for qPCR, detections must occur with value(s) above the LOD. o "Inconclusive" is defined as a detection in only 1 of the station or PCR technical replicates, and for qPCR, the detection must occur with a value above the LOD. These data will be permissible to submit with your dataset but will not be displayed on the map viewer. They will be retrievable in the table format with the rest of your submitted data.
Copies /L or /g (Above LOQ)	If quantitative detection was employed, enter the DNA copy number calculated per liter or gram (from water or sediment, respectively). For qPCR, this value must be above a calculated Limit of Quantification (LOQ) [See Box 2 above].

The following section provides supplemental information.

FREQUENTLY ASKED QUESTIONS (FAQ)

USGS NAS eDNA Webinar Frequently Asked Questions (FAQ) adapted from town hall presentations (originally prepared June 2020)

Concerns on displaying eDNA results

1) The interpretation of results from eDNA detections comes with many nuances, and a lack of public understanding of the science. Because the NAS Database is publicly available, there is a concern that many of the users will not understand what a "positive" eDNA result means. eDNA is a rapidly developing science, how will NAS data stewards ensure that the data will be put in a proper context and not misinterpreted by those without an understanding of the science?

The NAS Database Program is equally concerned that users have a clear and correct understanding and interpretation of eDNA results. To address this, plans include: 1) a short "Quick-start" summary document will be available on the website. The document will outline the basic science and definitions (to include detections/non-detections), the scope of the eDNA data layer, and attempt to provide context between the visual sighting data and eDNA detection data layers; 2) a disclaimer that appears when the eDNA data layer is activated in the map viewer where the user will be required to acknowledge having read and understood the meaning of a positive eDNA result; 3) tool tips (represented by question mark icons) will be available next to the layer in the legend and will provide helpful information about a data layer; and 4) references to journal articles to help users understand eDNA results.

2) What about taxa that are not currently included on the NAS Database – will you also be including data from them?

Inclusion of new species depends on the level of support by an agency or institution. The NAS Database data stewards are willing to consider adding a species not currently tracked by the NAS Database. It is always the NAS Database's goal to have a complete record of a species' invasion. We prefer not to include species that we have little expertise or support in aggregating new occurrence data into the future. Any data provided to the NAS Database but not added to the Database would be shared with the appropriate state jurisdiction, as allowed. Cross walking information to other databases focused on more broad species records is being explored.

3) Some scientists may feel uncertain in sharing a positive eDNA detection result without a live specimen associated with it; how do NAS data managers plan to keep stakeholders informed about the nuisances for eDNA data display in their areas of concern?

The NAS database is taking a conservative approach to sharing and interpretation of eDNA positive detection results. We are also working to ensure that data to be published on the NAS Database are from studies that have employed best practices and have run appropriate controls. No eDNA detection data will be displayed within a jurisdiction without the agreement/consent of the appropriate stakeholder agency when applicable (some agencies, such as the USGS, have policies against the sequential release of data).

Stakeholder involvement

4) The standards the NAS Database gives for the release of a record is that the stakeholder(s) will review and approve a record; what agencies and/or groups are considered stakeholders and will be making this decision?

A stakeholder is defined as an agency that has natural resource management authority within the jurisdiction where the detection was made. Within a State, this would include the lead agency of the State aquatic invasive species (AIS) coordinator, or on tribal lands it would be the First Nations' management authority. For boundary waters between two States, each State's AIS coordinator would be included and any other concerning authority. For federal lands, this would include the federal management agency and the appropriate state AIS coordinator. For certain data, sequential release of data to a stakeholder or manager is not allowed, and such data would be made available through the website to everyone at once.

5) Do you need consensus among all management groups to display eDNA data on the NAS Database website?

It is the NAS Database's goal to have a consensus among all management groups before displaying eDNA data in their jurisdiction (as allowed). This will provide all groups involved assurance that eDNA data will not be displayed without their knowledge. The NAS Database will provide options for the length of time data can be held to allow agencies time for a response. If a situation arose that groups did not consent for eDNA data to be displayed, the NAS Database and/or other outside groups would hold a series of meetings to help mediate any issues that have occurred.

6) What will be the amount of work required by stakeholders and NAS Database data stewards to meet the requirements to display eDNA, especially with all the data standards and verification form?

We have developed, with community input and consensus, a pre-submission Application to ensure studies followed best practices and ran appropriate controls. The standards we require for data submission may be considered rather stringent but are expected to evolve over time depending on the establishment of new standards or best practices in the eDNA field of science.

In order to submit data, questions on the survey must be answered in the affirmative, indicating that the study included the required controls. Many of these requirements are a result of an effort to meet consensus among a broad group of eDNA researchers and natural resource managers. As these requirements are based on best practices from the literature, it is assumed many studies already have met them. If approved, the NAS Database can upload thousands of eDNA detection/non-detection records in tabular format (e.g., .csv file) into the database.

7) Will the amount of work required by NAS Database data stewards to display eDNA change the efficiency of the program to perform its current duties?

We anticipate this effort will require a minimal amount of additional work to develop and display the eDNA detection data on the NAS Database. All the appropriate server and database infrastructure is in place. At this time, we do not foresee this resulting in significant impacts to our current work or overall program.

Displaying eDNA occurrences

8) What does it mean to be a verified data source? How do you become a verified data source?

A verified status is achieved after submitting a pre-submission Application for a study wherein the required questions are all answered in the affirmative, indicating the necessary best practices and controls were employed in the study. For such ongoing studies, where new data are submitted annually, the original pre-submission Application that was approved gives that individual or organization a verified status. They will not need to re-apply each time to submit more data so long as the new data was collected using the same methodology and their answers to the Application have not changed.

9) Will the NAS Database display detections at both the point and Hydrologic Unit Code (HUC) level?

The preference for any data in the NAS Database is precise specific geographic locations. This is not always feasible, or practical, based on the sampling methodology, so both point-level and fine scale HUC- level (twelve-digit hydrologic unit code) information will be available.

10) Will eDNA occurrences be shown on the HUC8 (species profile) sub-basin maps on the NAS Database?

The HUC8 sub-basin map found on each species profile will include information about positive eDNA data. The HUC8s with only eDNA data will have a different color to make them distinct from specimen- derived sub-basins.

11) Is there a way for a general user to be able to understand the meaning of high-low frequency when interpreting the eDNA data?

We will provide an interpretive guide describing any metric that is being used to visualize eDNA detection data. This guide will be accessible from the eDNA data layer on the NAS Database distribution point map.

12) Will USGS only be including data from the United States, or will you also include eDNA data from shared boundary waters and connected waterways with Canada and Mexico?

Initially we only plan to include data from waters of the United States. However, in the future we anticipate including eDNA data submitted by partners in neighboring countries and from bi-national agencies or organizations (i.e. Great Lakes Fishery Commission) provided they meet our community standards.

eDNA community standards

13) Will the NAS Database accommodate a qPCR assay that indicates more than one species?

Currently, the pre-submission Application requires assays be validated as being species specific. The approval process is purposefully rigid in this case. If a dataset is submitted from a qPCR assay suggesting multiple species (and the assay is not designed as a multiplexed qPCR), the data would likely be inadmissible. For applicants with data from multiple species, such as studies where separate individual assays are used to assess the same water or sediment samples, each dataset will require a separate application.

14) Will USGS define a standard of what a positive eDNA detection is (i.e. 1 in 3 vs. 3 in 3 technical replicates)?

Our approval process requires that a "Limit of Detection" is calculated based on 95% reproducibility, including requirements for blanks and controls. We do require at least 3 technical replications are run for each sample, and a "detection" requires more than one technical replicate be positive. In the case of this question, 2 of 3 technical replicates would be reportable and displayed as a detection. In cases where 1 of 3 are a positive detection, the sample may be reported as an "inconclusive" detection. This would not be displayed on the map viewer as a positive detection but would be available for review by someone who requests the entire dataset.

Null or non-detection eDNA data

15) Will the NAS Database display negative or non-detections? How will they be displayed on a map?

The NAS Database will display both positive detections and non-detections. In the case of eDNA surveys, a "non-detection" is the reporting of the specific analysis and does not preclude the chance that the eDNA was not somewhere in the environment and simply did not get collected into the sample bottle or container. Because of this possibility, the term "negative" detection is not preferred. The NAS Database has a distinct color scheme that clearly distinguishes non-detections (null) from positive results. It is our intention to explicitly identify the status of all eDNA results. The NAS species distribution maps have a legend to help identify elements on the map.

16) Can you provide a clear definition differentiating a null detection vs. a negative detection vs. a non-detection?

The eDNA community is moving towards "detect" and "non-detect." Null detection, negative detection, and non-detection are being used interchangeably in some cases. It is a matter of which term is more accepted. In each case, the terms indicate that an analysis of a sample did not detect the genetic material of a target species. We have chosen "non-detection" because it appears to be most preferred by the eDNA community (See Q.15 above).

17) The NAS Database specimen records are for presence-only of animals and plants, while the eDNA data displayed will be both presence and absence points; how does NAS intend to reconcile this discrepancy in the observations?

The traditional specimen records of visual observations in the NAS Database are generally limited to presence-only due to the opportunistic nature of most observations of introduced aquatic species. Very few types of traditional physical sampling gear and methodologies can be taxon specific. The NAS Database does have the capability to hold and display absence records from taxon-specific targeted surveys. However, we generally do not receive absence data from any data provider. eDNA assays are a type of targeted, taxon-specific sampling effort, and thus allows a high degree of confidence in both detections and lack of detections.

Historical eDNA data

18) What value would an eDNA sample collected prior to eradication of a population hold?

Pre-sampling gives a baseline for detection and occurrence rates, exhibits the success of the assay, and it helps to calibrate inputs for that assay.

19) Will you accept and display historical (not current) eDNA data?

If the data meet the community standards, we will happily accept and display historical data on the NAS Database website to provide a more complete picture of an invasion.

Metadata

20) What metadata will be included with eDNA data hosted on the NAS website?

The plan is for the metadata to be generated for each eDNA record and include pertinent spatio-temporal values, data fields describing information about the sampling, processing, and analysis of the samples, and some covariate information. The metadata will be available after peer review [For the final product, this "review" was achieved through the application itself and the metadata are reported in the data template].

21) Will the metadata contain a ranking that represents the likelihood of a detection being a false positive?

Not specifically, the only ranking we are considering for the future is to assign a value based off the overall study's relative adherence to best practices in addition to the required standards for submission. However, the pertinent metadata to analyze a dataset in a manner that allows one to calculate the likelihood of detection will be requested from the data submitter, though not required.

22) Will water quality parameters (pH, temperature, etc.) be included in the metadata?

A part of our current effort is the preparation of the data submission template. This template will contain data fields representing the geospatial and species data required for producing the data layer and data fields which capture methodological variables including sample replicates and the like. Covariates such as water temperature, pH, etc., were not identified as "required" metadata for verification of data prior to submission. However, future efforts to increase the metadata associated with submitted data to include such covariates is planned.

23) Will the NAS metadata record include information about level of uncertainty of the data?

Our effort was focused on a level of confidence based on the controls and best practices employed. We will not be assigning a value to the "level of uncertainty" of other researchers' results just as a journal would not when publishing papers. We are vetting the whole of the data during the submission process, and not the results directly.

eDNA communication plan

24) Will NAS Database develop communication plans with individual states?

An overarching communication plan has been developed for the application and submission process. However, it may be replaced by more detailed communication plans developed in concert with individual States, First Nations, and Federal agencies on a case-by-case basis. (See Ferrante, et. al, 2022).

Sharing eDNA data

25) When will the NAS Database be accepting data?

The NAS Database is planning on only accepting processed data from data donors. It is the NAS Database's goal to begin accepting eDNA data by fall of 2023.

26) Will there be an application programming interface (API) for the data?

There are plans to develop an API for eDNA data soon. In the short term, all eDNA data hosted on the NAS Database website would be available for download by anyone as is already the case for visual observation data.

GLOSSARY

A

Amplicon: product of amplification (or replication events) of a targeted DNA sequence (See: PCR)

B

Batch: group of samples processed and/or analyzed at the same time such as all the samples on a PCR plate or strip, or the subset of water samples filtered in a singular effort.

Biosurveillance: *J. Darling suggestion*: monitoring of species over time and space when the presence or absence of those species potentially have direct bearing on management decision-making.

Biomonitoring/ biodiversity monitoring: monitoring of a species over time and/or place

C

Conventional/ end point PCR: (See: PCR (Polymerase chain reaction))

D

Decision science: "the collection of quantitative techniques used to inform **decision**-making at the individual and population levels." (Link to reference)

Detect(ion): (see also: Non-detect(ion)) signal in PCR analysis indicating presence of target DNA.

DNA isolation: separation of DNA from other molecules in a sample, such as proteins, lipids, and other organic material. Common methods for DNA isolation are organic isolations (phenol-chloroform-isoamyl; PCI), chelex isolations, and solid phase isolations. DNA isolations kits are widely available and vary their use in input sample type.

Digital PCR (dPCR): PCR assay in which the sample is partitioned into thousands of independent reactions that are each analyzed as end-point PCR reactions. The proportion of positive reactions can be directly related to the concentration of target in the original sample without need of a standard curve and with reduced susceptibility to biochemical inhibition. The two major platforms for dPCR are droplet-based and chip-based. Because we represent the Federal government and do not endorse one platform over the other, we utilize the generic dPCR acronym which applies to both types of digital PCR.

Dye-based: a method where a non-sequence-specific marker (*e.g.*, a dye such as SYBR Green I) intercalates to nucleic acids to generate a fluorescence signal that is used to report and quantify dsDNA amplification

E

Efficiency (PCR): ratio of the actual number of target molecules replicated during a PCR cycle relative to the expected number replicated. If all molecules are replicated, the efficiency would be 100%.

environmental DNA (eDNA): genetic material (DNA) that is present in an environmental medium, such as water, sediment, or air. Often, it derives from sources shed by organisms into their environment including sloughed skin, saliva, feces, gametes, detritus and decay, although entire microscopic organisms can also be a source of eDNA if collected during sampling.

End point PCR: (See: Polymerase chain reaction (PCR))

Extraction: lysis and/or breakdown of organic or inorganic material in which eDNA is contained allowing it to be separated from the medium in which it was originally found in the environment.

F

Field equipment blank: a "sample" acquired using the field sampling equipment prior to a field collection effort wherein DNA-free water or soil from an off-site location is collected using the sampling equipment. This is often performed at the field vehicle or in boat before launching and involves pouring DNA-free water or dispensing DNA-free sediment into a collection receptacle (e.g., bottle, tube, bag, etc.) in the field.

Filtration: passing water samples through a membrane that captures eDNA-containing particles and is suitable for DNA preservation and later recovery. The pore size and membrane material are major considerations in selecting a filtration strategy, as is the need for any pre-filtering to eliminate larger particles and sediments that would quickly clog filters with small pore sizes. Filtration can be integrated with the field sampling process, e.g., by using positive or negative pressure to pass water through a flow-through system, or alternatively deferred to a different time or setting if the sample integrity can be preserved.

G

H

Hydrolysis probe: Labeled probes used in quantitative PCR comprised of a primer, a fluorescent reporter, and often a quencher. They are to target specific genetic sequences of a

target species. When the probe aligns with the target sequence, fluorescence is emitted by the reporter and detected by the fluorescence reader in a qPCR or dPCR system.

I

Inconclusive detection: Target species PCR detection that either has not been reproduced or does not exceed the limit of detection (LOD)

Inhibition/Inhibitor: (CSAS def) - non-target substances from the environmental system (e.g., total suspended solids) that remain present in the sample at collection and through DNA extraction. Inhibitors are typically co-extracted with the target DNA and inhibit the PCR/qPCR reaction. Failure to test / characterize the presence of inhibition in samples can result in false negatives.

J

K

L

Limit of Detection (LOD): (G) The lower limit of DNA concentration in copies/L that a presence assay can robustly detect, estimated by serial dilution of a positive control sequence of known concentration. This control DNA could be from a size-selected and purified PCR product or plasmid-cloned PCR product, for example that has been confirmed by sequencing.

Limit of quantification (LOQ): from (Klymus et al. 2020), reflects the assay's capacity to precisely quantify copy number. LOQ is calculated as the lowest standard concentration that could be quantified with a Coefficient of Variation (CV) value below 35% where CV is calculated for each standard by the equation derived by Forootan et al. (Forootan et al. 2017).

M

Metabarcoding: a method used to identify all organisms in an eDNA sample by using general primer sets in PCR, usually with limits to taxa of interest.

N

No template control (NTC): alternatively named a negative control, this is a reaction where all reagents and no template (DNA, standard, etc.) is added to ensure no background signal is being reported/detected.

Non-detect(ion): the lack of amplification of target DNA indicating that target eDNA was not present in the sample/qPCR replicate **(see also: Detect(ion))**

O

Occurrence (rates): A quantitative calculation of the probability of the target species occurring at a location.

P

Polymerase chain reaction (PCR): (G) A process by which a segment of a continuous DNA molecule can be cyclically copied in a controlled reaction to create orders of magnitude more copies of that molecule. PCR is used to create a high concentration of uniform DNA fragments that can be easily visualized or used in other reactions such as a DNA sequencing reaction. The chain reaction is driven by cyclical changes in the temperature of the reaction, causing a DNA polymerase to repeatedly associate with the priming site and then disassociate from the completed copy when each cycle is complete. A high level of specificity is achieved by using "primer" sequences that must closely match the intended target sequence, and in the intended orientation, for efficient amplification to occur under the (optimized) reaction conditions.

Variants of a PCR assay include end-point PCR (or "conventional" PCR), in which the presence or absence of the intended target is inferred only after the reaction completes the amount of input target DNA is inferred by monitoring the change in DNA concentration as the reaction proceeds; and digital-droplet PCR, which also quantifies the amount of input target DNA by creating many random independent reactions from the input solution and determining the proportion of which produce an amplification product.

Probe: – (G) In the context of the "TaqMan" class of protocols for qPCR, a probe is an oligonucleotide that complements the intended target and is added along with primers to a qPCR reaction. The probe bears a fluorescing molecule on one end and a suppressor of fluorescence ("quencher") on the other end. Fluorescence only occurs when the probe anneals to a copy of the target sequence that is being actively replicated by a DNA polymerase enzyme. As the DNA polymerase progresses along the template and encounters a bound probe, the fluorophore is liberated from the quencher and is then detected by the sensor. Probes can increase the specificity of qPCR reactions because the detected signal comes only from DNA templates that match closely at both primer sites and the probe site.

Q

Quantitative PCR (qPCR): Quantitative or real-time PCR that is used to quantify the amplification of the target DNA while the reaction is being performed using a standard curve.

R

Region: Larger geographic units that encompass numerous sites – a park or preserve, drainage basin, county, etc. (compare to Site, Station)

S

Site: (CSAS def - are physical places where samples have been collected; sites should be relatively independent of each other, such as different systems and habitats (e.g., different lakes, rivers, ponds, marinas, tributaries of different order). (Compare to Region, Station)

Station: (CSAS def - refers to spatially distinct sampling locations within a site (i.e., spatial replicates) and are typically used to improve species detection or evaluate the eDNA variation within systems or habitats (e.g., samples distributed using a grid or transect design surrounding an aquaculture site, upper and lower reach of a river, locations within large open water environments). (Compare to Region, Site)

Synthetic eDNA: DNA molecules created in a laboratory for use in eDNA assays. They are commonly used as standards in qPCR and digital PCR

T

Technical replicates – the method of testing the sample multiple times to test the variability within the test itself. These are commonly used in quantitative PCR and digital PCR.

U

V

W

X

Y

Z

Placeholder terms (not yet in text)

Contaminated positive: a sample or control that has been contaminated, usually by technical or human error.

Presampling – environmental sampling done prior to the application of a treatment to understand the effects of the treatment

REFERENCES

Abbott C, Coulson M, Gagné N, Lacoursière-Roussel A, Parent GJ, Bajno R, Dietrich C and May-McNally S (2021) Guidance on the Use of Targeted Environmental DNA (eDNA) Analysis for the Management of Aquatic Invasive Species and Species at Risk. Fisheries and Oceans Canada, Ottawa, ON, Canada, Canadian Science Advisory Secretariat, 2021/019**:** iv+42 p., https://publications.gc.ca/site/eng/9.899437/publication.html.

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, et al (2009) "The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments." *Clinical Chemistry* 55 (4): 611–22. https://doi.org/10.1373/clinchem.2008.112797.

Dorazio, Robert M., and Richard A. Erickson (2018) "Ednaoccupancy: An R Package for Multiscale Occupancy Modelling of Environmental DNA Data." *Molecular Ecology Resources* 18 (2): 368–80. https://doi.org/10.1111/1755-0998.12735.

Dunker KJ, Sepulveda AJ, Massengill RL, Olsen JB, Russ OL, Wenburg JK and Antonovich A (2016) Potential of Environmental DNA to Evaluate Northern Pike (*Esox lucius*) Eradication Efforts: An Experimental Test and Case Study. *PLoS One* 11(9): e0162277, https://doi.org/10.1371/journal.pone.0162277

Ferrante JA, Daniel WM, Freedman JA, Klymus KE, Neilson ME, Passamaneck Y, Rees CB, Sepulveda A and Hunter ME (2022) Gaining decision-maker confidence through community consensus: identifying environmental DNA standards for data display on the USGS Nonindigenous Aquatic Species database. *Management of Biological Invasions* 13(4): 809-832, https://doi.org/https://doi.org/10.3391/mbi.2022.13.4.15

Forootan A, Sjöback R, Björkman J, Sjögreen B, Linz L and 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, https://doi.org/https://doi.org/10.1016/j.bdq.2017.04.001

Goldberg CS, Turner CR, Deiner K, Klymus KE, Thomsen PF, Murphy MA, Spear SF, McKee A, Oyler-McCance SJ, Cornman RS, Laramie MB, Mahon AR, Lance RF, Pilliod DS, Strickler KM, Waits LP, Fremier AK, Takahara T, Herder JE and Taberlet P (2016) Critical considerations for the application of environmental eDNA methods to detect aquatic species. *Methods in Ecology and Evolution* 7(11): 1299-1307, https://doi.org/10.1111/2041-210X.12595

Huggett JF, Foy CA, Benes V, Emslie K, Garson JA, Haynes R, Hellemans J, et al (2013) "The Digital MIQE Guidelines: Minimum Information for Publication of Quantitative Digital PCR Experiments." *Clinical Chemistry* 59 (6): 892–902. https://doi.org/10.1373/clinchem.2013.206375.

Hunter ME, Dorazio RM, Butterfield JSS, Meigs-Friend G, Nico LG and Ferrante JA (2017) Detection limits of quantitative and digital PCR assays and their influence in presence-absence surveys of environmental DNA. *Molecular Ecology Resources* 17(2): 221-229, https://doi.org/10.1111/1755-0998.12619

Hunter ME, Meigs-Friend G, Ferrante JA, and Ulmer A (2019) "Improving eDNA Yield and Inhibitor Reduction through Increased Water Volumes and Multi-Filter Isolation Techniques." *Scientific Reports* 9 (1): 5259. https://doi.org/10.1038/s41598-019-40977-w.

Hunter ME, Meigs-Friend G, Ferrante JA, Smith BJ and Hart KM (2019) Efficacy of eDNA as an early detection indicator for Burmese pythons in the ARM Loxahatchee National Wildlife Refuge in the greater Everglades ecosystem. *Ecological Indicators* 102: 617-622, https://doi.org/10.1016/j.ecolind.2019.02.058

Hunter ME, Oyler-McCance SJ, Dorazio RM, Fike JA, Smith BJ, Hunter CT, Reed RN and Hart KM (2015) Environmental DNA (eDNA) sampling improves occurrence and detection estimates of invasive Burmese pythons. *PLoS One* 10(4): e0121655, https://doi.org/10.1371/journal.pone.0121655

Klymus KE, Merkes CM, Allison MJ, Goldberg CS, Helbing CC, Hunter ME, Jackson CA, Lance RF, Mangan AM, Monroe EM, Piaggio AJ, Stokdyk JP, Wilson CC and Richter CA (2020) Reporting the limits of detection and quantification for environmental DNA assays. *Environmental DNA* 2(3): 271-282, https://doi.org/10.1002/edn3.29

Laramie MB, Pilliod DS and Goldberg CS (2015) Characterizing the distribution of an endangered salmonid using environmental DNA analysis. *Biological Conservation* 183: 29-37, https://doi.org/10.1016/j.biocon.2014.11.025

Rees HC, Maddison BC, Middleditch DJ, Patmore JRM, and Gough KC (2014) "Review: The Detection of Aquatic Animal Species Using Environmental DNA - a Review of eDNA as a Survey Tool in Ecology." Edited by Erika Crispo. *Journal of Applied Ecology* 51 (5): 1450–59. https://doi.org/10.1111/1365-2664.12306.

Strickler KM, Fremier AK, and Goldberg CS (2015) "Quantifying Effects of UV-B, Temperature, and PH on EDNA Degradation in Aquatic Microcosms." *Biological Conservation* 183 (March): 85–92. https://doi.org/10.1016/j.biocon.2014.11.038.

Thalinger B, Deiner K, Harper LR, Rees CH, Blackman RC, Sint D, Traugott M, Goldberg CS, and Bruce K (2020) "A Validation Scale to Determine the Readiness of Environmental DNA Assays for Routine Species Monitoring," April. https://doi.org/10.1101/2020.04.27.063990.

Wilcox TM, McKelvey KS, Young MK, Jane SF, Lowe WH, Whiteley AR and Schwartz MK (2013) Robust Detection of Rare Species Using Environmental DNA: The Importance of Primer Specificity. *PLoS One* 8(3): e59520, https://doi.org/10.1371/journal.pone.0059520