

RESEARCH ARTICLE

Simultaneous species detection and discovery with environmental DNA metabarcoding: A freshwater mollusk case study

Daniel D. Vanderpool  | Taylor M. Wilcox  | Michael K. Young  | Kristine L. Pilgrim | Michael K. Schwartz

USDA Forest Service Rocky Mountain Research Station, National Genomics Center for Wildlife and Fish Conservation, Missoula, Montana, USA

Correspondence

Daniel D. Vanderpool, USDA Forest Service Rocky Mountain Research Station, National Genomics Center for Wildlife and Fish Conservation, Missoula, MT, USA.
Email: daniel.vanderpool@usda.gov

Funding information

The Strategic Environmental Resource Development Program, Grant/Award Number: Project RC18-1348; The Rocky Mountain Research Station

Abstract

Environmental DNA (eDNA) sampling is a powerful tool for rapidly characterizing biodiversity patterns for specious, cryptic taxa with incomplete taxonomies. One such group that are also of high conservation concern are North American freshwater gastropods. In particular, springsnails of the genus *Pyrgulopsis* (Family: Hydrobiidae) are prevalent throughout the western United States where >140 species have been described. Many of the described species are narrow endemics known from a single spring or locality, and it is believed that there are likely many additional species which have yet to be described. The distribution of these species across the landscape is of interest because habitat loss and degradation, climate change, groundwater mining, and pollution have resulted in springsnail imperilment rates as high as 92%. Determining distributions with conventional sampling methods is limited by the fact that these snails are often <5 mm in length with few distinguishing morphological characters, making them both difficult to detect and to identify. We developed an eDNA metabarcoding protocol that is both inexpensive and capable of rapid, accurate detection of all known *Pyrgulopsis* species. When compared with conventional collection techniques, our pipeline consistently resulted in detection at sites previously known to contain *Pyrgulopsis* springsnails and at a cost per site that is likely to be substantially less than the conventional sampling and individual barcoding that has been done historically. Additionally, because our method uses eDNA extracted from filtered water, it is non-destructive and suitable for the detection of endangered species where “no take” restrictions may be in effect. This effort represents both a tool which is immediately applicable to taxa of high conservation concern across western North America and a case study in the broader application of eDNA sampling for landscape assessments of cryptic taxa of conservation concern.

KEYWORDS

environmental DNA species detection, metabarcoding, springsnails

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Published 2024. This article is a U.S. Government work and is in the public domain in the USA. *Ecology and Evolution* published by John Wiley & Sons Ltd.

TAXONOMY CLASSIFICATION

Conservation genetics, Environmental sustainability

1 | INTRODUCTION

Environmental DNA sampling—the inferences of species presence from genetic material in the environment—is unique among sampling methods in that there is the opportunity to characterize genetic diversity within entire ecological communities, capturing information about described and undescribed taxa (Deiner et al., 2017). We can simultaneously determine the occupancy status for known species of concern and also provide new information about the existence of previously unknown genetic lineages on the landscape (e.g., Clusa et al., 2016). This use of eDNA sampling is particularly relevant for taxonomic groups of conservation concern for which current taxonomies are incomplete and morphological identification is difficult.

One example of a cryptic and taxonomically incomplete group of conservation concern are the freshwater snails of North America. Freshwater snails have limited dispersal abilities, narrow environmental niches, and often persist in geographically isolated habitats—all of which contribute to high levels of endemism and cryptic speciation (Hershler, Liu, & Howard, 2014; Johnson et al., 2013; Stevens et al., 2022; Young et al., 2021). This group is a focus of conservation due to pervasive habitat degradation of their native waters and competition with nonnative species invasions (Hershler, Liu, & Howard, 2014; Johnson et al., 2013). The conservation status of the most species-rich family of freshwater snails, Hydrobiidae, is of particular concern. Of the 185 described Hydrobiidae species in North America, 16 genera and 170 species are regarded as imperiled and 14 species are listed under the U.S. Endangered Species Act (Johnson et al., 2013). Most taxa ($n=139$) within this family are in the genus *Pyrgulopsis*. They are primarily found in the southwestern United States and often inhabit isolated perennial springs, from which they derive their common name, springsnails (Hershler, Liu, & Howard, 2014). This genus is of even greater concern because it is known to contain cryptic taxa that have yet to be described and additional lineages that have yet to be discovered (Hershler & Liu, 2017; Hershler, Liu, & Howard, 2014; Hershler, Ratcliffe, et al., 2014; Liu et al., 2017).

Environmental DNA sampling is a powerful tool for detecting freshwater snails—and *Pyrgulopsis* springsnails in particular—for several reasons. First, across much of their range, occupancy status in potential habitats has yet to be determined or updated for *Pyrgulopsis* (Hershler, Liu, & Howard, 2014). Thus, rapid eDNA sampling protocols will allow large landscapes to be comprehensively and frequently inventoried by crews with limited training. Second, morphological identification of *Pyrgulopsis* taxa is difficult, such that most current sampling for this group is reliant on destructive genetic analyses (Liu et al., 2003; McKelvey et al., 2020). This is both expensive and potentially problematic as some species are

federally protected (e.g., Proposed Rule 80 Federal Register 2015-25058). Similarly, eDNA sampling creates the potential to sample haplotypes from across an entire population of springsnails (or at least more than the 10–30 individuals typically collected for destructive analysis), which enables greater power to detect low-frequency haplotypes. In the case of newly discovered species, conventional sampling will still be required to perform morphological descriptions for species designation, but these sites likely represent a minority of cases.

The high levels of genetic diversity among *Pyrgulopsis* mitochondrial haplotypes (Hurt, 2004; Lavretsky et al., 2021; Liu et al., 2017) also presents some possible challenges for taxon-specific eDNA sampling. Although taxon-specific qPCR is a sensitive and cost-effective tool for species detection, the high levels of endemism and the prevalence of many undescribed species within this group (Hershler, 1998; Hershler, Liu, & Howard, 2014) preclude a taxon-specific approach. Instead, we must rely on metabarcoding/amplicon sequencing-based approaches that allow for a broader taxonomic scope. Executing this effectively requires careful attention to primer development because current amplicon sequencing protocols (i.e., metabarcoding) can be highly sensitive to amplification bias (Kelly et al., 2019).

In this study, we use the genus *Pyrgulopsis* as a case study to develop and demonstrate a pipeline for species detection and discovery based on eDNA metabarcoding. We build group-specific primers by assembling a cytochrome oxidase subunit I (COI) sequence alignment and phylogenetic tree for *Pyrgulopsis* taxa in North America and target a taxonomically informative portion of the gene. We then collected eDNA samples from across the southwestern United States and compared the effectiveness of this approach for population detection with conventional sampling as part of a multi-year monitoring program in a portion of the study area.

2 | METHODS

2.1 | Primer and positive control development

To develop degenerate primers to preferentially amplify *Pyrgulopsis* species, we first ensured that we had representative sequences spanning the breadth of *Pyrgulopsis* diversity. To accomplish this, we constructed a phylogenetic tree for *Pyrgulopsis* using COI sequences available from GenBank as of June 2023. After filtering the dataset to exclude multiple representatives from the same species complex, we constructed an alignment of 137 taxa (131 *Pyrgulopsis*, six outgroup taxa) using MAFFT v7.215 (Katoh & Standley, 2013), including the majority of described

Pyrgulopsis species. The resulting 658 base pair (bp) alignment was partitioned by codon position into three ~220bp character sets. We fit nucleotide substitution models to each character set using ModelFinder (set to merge partitions for a better fit) as implemented in IQTREE 2.1.3 followed by maximum likelihood tree inference with 1000 ultra-fast bootstrap replicates (Chernomor et al., 2016; Hoang et al., 2018; Kalyaanamoorthy et al., 2017; Lanfear et al., 2012; Minh et al., 2020). Limited phylogenetic signal for resolving deeper nodes within the group led us to perform 15 independent searches in order to ensure convergence of likelihood values among multiple runs (command line: `iqtree2 -s Alignment_file.fa -p partition_file.nex -runs 15 -m MFP+MERGE -B 1000 -nt 2`). The resulting maximum likelihood topology revealed a monophyletic *Pyrgulopsis*, with four weakly supported primary clades. Since the included *Pyrgulopsis* species formed a monophyletic clade in our phylogenetic analysis we elected to use this taxon set for downstream analyses and primer construction (The alignment file, codon partition file, and resulting tree are available on Dryad <https://doi.org/10.5061/dryad.tb2rbp07q>, Vanderpool et al., 2023a).

We designed degenerate primers capable of amplifying the 3' region of the *Pyrgulopsis* mitochondrial gene COI. Similar to other universal COI primers used in metabarcoding (Elbrecht & Leese, 2017; Folmer et al., 1994; Geller et al., 2013), the reverse primer was designed to exploit a conserved region near the 3' end of the gene. To increase specificity for *Pyrgulopsis* we reduced the degeneracy from that of the BR2 primer described by Elbrecht and Leese (2017); 144-fold degenerate versus 196-fold degenerate in the original primer. However, since all but one of the 131 *Pyrgulopsis* COI sequences obtained from GenBank were amplified and sequenced using this same reverse priming site, the priming site sequence was not available for alignment. Instead, we performed BLAST searches with *Pyrgulopsis* COI queries to identify the most similar gastropod species where the priming site sequence was available. We then aligned the gastropod COI sequences identified with this search to our *Pyrgulopsis* representatives and used the states present in these taxa to reduce the degeneracy in the reverse primer. The forward primer is a version of BF1 described by Elbrecht and Leese (2017), modified to decrease degeneracy (96-fold degenerate versus 128-fold degenerate) according to our alignment of 131 *Pyrgulopsis* COI sequences described above.

The true number of *Pyrgulopsis* species is unknown, making it difficult to ascertain whether our primers with decreased degeneracy may fail to detect some new species (Eloe-Fadrosh et al., 2016). The forward priming site is invariant at the first and second codon positions; however, all six of the third codon positions are 2- to 3-fold degenerate in our alignment. To evaluate the potential for failing to detect undescribed *Pyrgulopsis* taxa, we constructed a test dataset containing hypothetical sequences with 1–3 primer-template base pair mismatches in the forward primer, reverse primer, or both. We then conducted an in silico test of the effects of potential mismatches on our ability to amplify *Pyrgulopsis* DNA using eDNAAssay (Kronenberg et al., 2022), a random forest classifier, to predict

amplification of these hypothetical templates with up to three base pair mismatches at each primer site.

The primers described above were tested for effectiveness using pure extractions of DNA from three species: *P. ignota*, *P. metcalfi*, *P. texana*. Because these amplifications were successful (data not shown), we tested these primers using known positive eDNA extracts (based on 16S sequencing), as well as extracts thought to be negative for *Pyrgulopsis* DNA to gauge the effectiveness of amplifying target species in a mixed extract. We added an Illumina adapter sequence to our primers for the second indexing-PCR step that includes a unique barcode identifier for each sample.

Pyrg_BF1_tailed-GTGACTGGAGTTCAGACGTGTGCTCTTCCG
ATCTACBGGRTGRACYGTRTAYCC

Pyrg_BR2_tailed-ACACTCTTCCCTACACGACGCTCTTCCGAT
CTTCDGGRTGHCCRAARAAYCA

It is common for eDNA extractions to contain PCR inhibitors, motivating the addition of an inhibitor removal step to our protocol (McKee et al., 2015). However, even with this additional step, the possibility of residual inhibitors can result in a false negative test. In order to detect possible inhibition, we developed a positive PCR control using synthetic mitochondrial sequence from the extinct thylacine (Tasmanian Tiger, *Thylacinus cynocephalus*). Priming sites consisting of *Pyrgulopsis* sequences compatible with our degenerate primers were added to the end of the thylacine sequence, resulting in an amplification product 100bp shorter than the amplicon created from true *Pyrgulopsis* templates. The use of thylacine DNA makes it possible to distinguish potential contamination from the positive control from that of background contamination derived from extant species present in the lab space (very few molecular laboratories are actively working on museum specimens of thylacines). The shorter amplicon size makes it possible to spike a small amount of positive control into one of the sample PCR replicates to test for reaction inhibition but differentiate the thylacine control from *Pyrgulopsis* targets on an agarose gel. We assumed there was PCR inhibition if all the reactions, including those with the positive control, did not amplify. If only the reaction containing the positive control amplified, and the product was 100bp shorter than expected, we assumed the sample was not inhibited, but that it was negative for target species DNA. When all or some of the replicates amplified, the one containing the positive control had a double band, was free from inhibitors, and potentially contained amplicons for the target species. We then selected potentially positive samples for indexing PCR and sequencing.

For this project, the linear, synthetic gene for the thylacine was purchased from Integrated DNA Technologies (gBlock; IDT), resuspended in Tris-EDTA, and quantified on a Qubit fluorometer (ThermoFisher Scientific) prior to dilution to the target copy number concentration. The positive control sequence with priming sites underlined is: 5'-ACTGGGTGGACCGTGTATCCTCATCACAACAATTATCAATATAAAACCTCCAGCCCTATCCCAATATCAAACTCCATTATTGTTTGATCAGTAATAATTACAGCAGTACTTCTACTATTATCCTTACCCGTCTGGCAGCAGGCATTACAATACTACTTA

CGGACCGAAATCTTAATACAACATTTTTTGACCCTGCTGGAGG
AGGTGACCCAATCCTTTATCAACATCTATTTTGATTTTTTGGTC
ACCCTGA-3'.

2.2 | Sampling and analysis

Our goal was to characterize springsnail biodiversity focusing on habitats with few or no historical data on these taxa. We conducted eDNA sampling in potential springsnail habitats, based on previous spring surveys or satellite imagery, on or near five U.S. Department of Defense installations in the southwestern U.S.: Fallon Range Training Complex (60 samples from 60 sites; Pershing and Churchill Counties, Nevada), Fort Huachuca (52 samples from 44 sites; Santa Cruz County, Arizona; including locations on Coronado National Forest and The Nature Conservancy Ramsey Canyon Reserve), Camp Navajo (4 samples from 4 sites; Coconino County, Arizona), Nellis Air Force Base (22 samples from 21 sites; Nye and Clark County, Nevada), and Camp Williams (2 samples from 2 sites; Salt Lake County, Utah). We also collected two additional samples on Bureau of Land Management lands in Utah (Figure 2). Samples were collected by U.S. Forest Service, Springs Stewardship Institute, Arizona Game and Fish Department, and Department of Defense crews. All sampling on Department of Defense lands was performed by either local installation personnel or sampling crews with research sampling permits and permissions.

We collected water samples following the protocol described by Carim, Dysthe, et al. (2016), Carim, McKelvey, et al. (2016) and Wilcox et al. (2016). Briefly, at each site we filtered up to 5 L of water through one to three 47 mm diameter, 1.5 μ m pore-size glass filters (GE Healthcare; www.gehealthcare.com) using a 12 V peristaltic pump. At each site, our goal was to filter 5 L of water. If the first filter paper clogged before reaching this goal, we continued filtering with a second filter. We continued with additional filter papers until we reached 5 L of total water or three filter papers. When necessary, multiple filter papers per site were combined during DNA extraction (below). We removed each filter with sterile forceps, folded it into quarters, and placed the filter in a sterile bag containing silica desiccant. For each filter paper collection, we used new, clean collection materials including the filter funnel and forceps. All other collection apparatus, including suction hoses and buckets, were washed with a 10% household bleach solution (approximately 0.6% sodium hypochlorite) to prevent cross contamination and inadvertent species transfer between sites. The desiccated samples were maintained at ambient temperatures for no more than 2 weeks before being transferred to a -20°C freezer prior to DNA extraction.

Extractions were performed in a dedicated eDNA extraction room kept free of PCR products and other potential background contamination in accordance with the National Genomics Center for Wildlife and Fish Conservation eDNA Program Standard Operating Procedure. We cut each filter in half; one half of the filter was archived at -20°C for future analyses and the other half was extracted

using a modified DNeasy Blood and Tissue Kit (QIAGEN) protocol. We followed the manufacturer's protocol for extraction from animal tissue with the modifications described in Carim, Dysthe, et al. (2016), Carim, McKelvey, et al. (2016). In cases where a sample consisted of multiple filters, we extracted each filter separately and then spun them through the same silica spin column (see Carim, Dysthe, et al., 2016). Extractions were eluted in 100 μ L of Tris-EDTA and then passed through a OneStep™ PCR Inhibitor Removal Kit column (Zymo Research) and stored at -20 or -80°C until PCR amplification.

PCR reactions were composed of 500 nM each primer (forward and reverse), 2 mM MgCl_2 , 2 ng BSA, 12.5 μ L KAPA HiFi 2 $\times$ HotStart ReadyMix (Roche, Inc.), 4 μ L template, and molecular grade water to a total reaction volume of 25 μ L. Reactions were run in quadruplicate with 1 μ L of the positive control added to one reaction in order to test for PCR inhibition. Thermal cycling parameters were designed to favor primer-template exact matches using a touchdown protocol. Samples were subjected to an initial denaturation at 95°C for 3 min, followed by 98°C for 20 s and then annealing temperatures starting at 67°C for 60 s, followed by a 20 s extension step at 72°C . The annealing temperature was decreased by one degree each cycle for 10 cycles until the optimal annealing temperature of 57°C was reached. This touchdown phase was followed by 40 cycles at 57°C with a 3-min extension step on the final cycle. The resulting PCR products were visualized on a 1% agarose gel, with successful reactions combined, cleaned using 0.9 $\times$ Ampure XP beads (Beckman Coulter), and eluted in 20 μ L of 10 mM Tris-EDTA. Illumina sequencing adaptors and barcodes were then added to each cleaned amplicon reaction via a 15 cycle indexing PCR consisting of the following reaction components: 3 μ L cleaned amplicon PCR product, 500 nM Illumina P5 adaptor with unique barcode, 500 nM Illumina P7 adaptor with unique barcode, 7 μ L H_2O , and 12.5 μ L KAPA HiFi 2 $\times$ HotStart ReadyMix (Roche, Inc.) to a final reaction volume of 25 μ L. Cycling parameters consisted of an initial denaturation of 98°C for 30 s, followed by 98°C for 10 s, 60°C for 20 s, and 72°C for 20 s. Indexed PCR products were cleaned using 0.9 $\times$ Ampure XP beads and eluted in 20 μ L of 10 mM Tris-EDTA. The concentration of each cleaned product was estimated using a Qubit fluorometer. Thirty to forty indexed samples were then pooled in equimolar ratios (~ 80 ng of each indexed sample) with the final volume cleaned using 0.9 $\times$ Ampure XP beads. This third cleaning step is critical for the complete removal of primer dimers generated during the preceding PCR steps. The concentration of the cleaned, pooled library was estimated using both a Qubit fluorometer and an Agilent 2100 TapeStation.

Pooled libraries were sequenced using a 2 $\times$ 250 cycle Illumina MiSeq Nano kit on four separate sequencing runs. However, one of these runs resulted in low yield and was sequenced a second time. For the second sequencing attempt, we selected seven sites and prepared two new indexed libraries from the same extractions as the first sequencing attempt. We sequenced these two libraries independently to verify the results of the initial low-yield sequencing run. Finally, we prepared two libraries for two of our samples for which we knew *Pyrgulopsis* were present in order to evaluate reproducibility in the detected haplotypes.

We processed reads using a pipeline based on the DADA2 package (Callahan et al., 2016) in R (R Core Team, 2020). We used the DADA2 function `filterAndTrim` to remove sequences with ambiguous bases or more than two expected errors based on quality score estimates, to trim off primer sequences, truncate sequences at the first quality score of four or to a maximum length of 250bp, and remove any sequences <25bp after trimming. We then used DADA2 to estimate error rates from our sequences using pseudo-pooled sampling, inferred amplicon sequence variants (ASVs), merged paired-end reads, and removed chimeric sequences. Next, we performed a BLAST search (Camacho et al., 2009) with *Pyrgulopsis marcida* COI sequence as a query to identify gastropod sequences close to *Pyrgulopsis* that were retained after filtering (command line: `blastn -query ../Pyrg_marcida_COI.fa -subject All_MiSeq_combined_DADA2_filtered.fa -evalue 1e-50 -outfmt 6`). The resulting sequences were subjected to further downstream inspection to eliminate any remaining errors. Three low-abundance ASVs were divergent from more common haplotypes found at the same site (typically <1% divergent from one another). The three ASVs differed in the ~60bp near the 5' or 3' end but were otherwise identical to the common haplotype. A BLAST search using this divergent region suggested that these were chimeric sequences, as the closest hit in each case was a species common in freshwater environments. Consequently, we inspected the remaining ASVs approximately 5% divergent from a common ASV and discovered another chimeric sequence with moderate abundance (322 reads) which was removed from downstream analyses. Additionally, in two instances a haplotype identified as unique was the reverse complement of the region of interest. These sequences were collapsed into their respective haplotypes. To identify potential contaminants, we explored detections with at or near 100% matches to a known sequence, but where the known distribution of the species of interest was geographically incongruent with the site in which the detection was made. Additionally, we used BLAST to compare *Pyrgulopsis* ASVs identified in our sequencing effort to a comprehensive FASTA file of springsnails previously sequenced for COI at the National Genomics Center for Wildlife and Fish Conservation ($n=2955$).

To assess the taxonomic affiliation of the ASVs retained after filtering, we added them to our original alignment, then estimated a phylogenetic tree using IQ-Tree 2.1.3 with 15 independent runs using the `--runs` flag implemented in IQ-Tree and retained the topology with the highest likelihood (command line: `iqtree2 -s Alignment_file.fa -p partition_file.nex --prefix Name_of_run --runs 15 -o Outgroup_name -m MFP -B 1000 -nt 2`). The majority of taxa were then trimmed from the resulting 207 taxon tree for visualization purposes (Figure 1).

2.3 | Comparison with conventional sampling

We compared results of our eDNA sampling and analysis to data from conventional sampling at 18 sites to determine concordance between *Pyrgulopsis* detection methods. We collected 1–5 eDNA

samples in either 2019 or 2020 from these 18 sites. Conventional samples were collected by a long-term *Pyrgulopsis* monitoring program in the Fort Huachuca area (Tullo & Sorensen, 2021). This program sampled the 18 sites for *Pyrgulopsis* detection 1–4 times from 2016 to 2020. Conventional surveys were conducted using active, timed visual searches of springsnails on natural substrates, typically lasting 10 to 20min or until first springsnail detection. All observed *Pyrgulopsis* were considered to be Huachuca springsnail and did not undergo further morphological or molecular analysis. Additionally, conventional sampling detections were sometimes variable over time and it is unclear to what extent this variation is due to imperfect detection versus year-to-year extirpations and recolonizations of local *Pyrgulopsis* populations. To attempt to control for this, we considered any site with at least one detection in 5 years to be a positive detection for our methods comparison.

3 | RESULTS

In addition to testing primer performance on DNA from several *Pyrgulopsis* taxa, we evaluated primer generality against potential, unseen diversity in *Pyrgulopsis* using an in silico tool (eDNAAssay; Kronenberger et al., 2022) and hypothetical templates with introduced mismatches (see Section 2). Even with up to four mismatches (two per primer), assignment probabilities were ≥ 0.76 . With up to three mismatches in a single primer, eDNAAssay assignment probabilities were ≥ 0.81 . Kronenberger et al. (2022) found that 100% of primer sets with assignment probabilities >0.7 amplified. This is an unrealistically large number of mismatches as the first and second codon positions located in the priming sites are fixed across the *Pyrgulopsis* examined here as well as highly conserved among divergent outgroups. The generality of these primers was further supported by the recovery of ASVs from divergent gastropods from environmental samples for which the nearest BLAST hits were to genera outside of Hydrobiidae, (e.g., *Turbo*, *Valvata*, *Glyptophysa*, *Rabdotus*, *Biomphalaria*, *Illana*, and *Muticaria*). None of the divergent ASVs recovered were close matches to these genera, likely meaning they were derived from species not represented in GenBank. Across these in silico and in vitro tests, we found strong evidence that the primer sets and PCR conditions used in this study were sufficient to sample all previously described *Pyrgulopsis* taxa and the range of variability that could be reasonably expected in as-of-yet undescribed taxa from this group.

3.1 | Sequencing and quality control

Our sequencing effort resulted in 53,894,920 paired-end reads (Vanderpool et al., 2023b). After filtering, denoising, and removing chimeric sequences using the DADA2 pipeline, these were pared to 30,280,265 reads across 56,171 ASVs. The majority of these reads were not *Pyrgulopsis* in origin (97.71%). Using BLAST to identify *Pyrgulopsis* haplotypes resulted in 127 ASVs, from

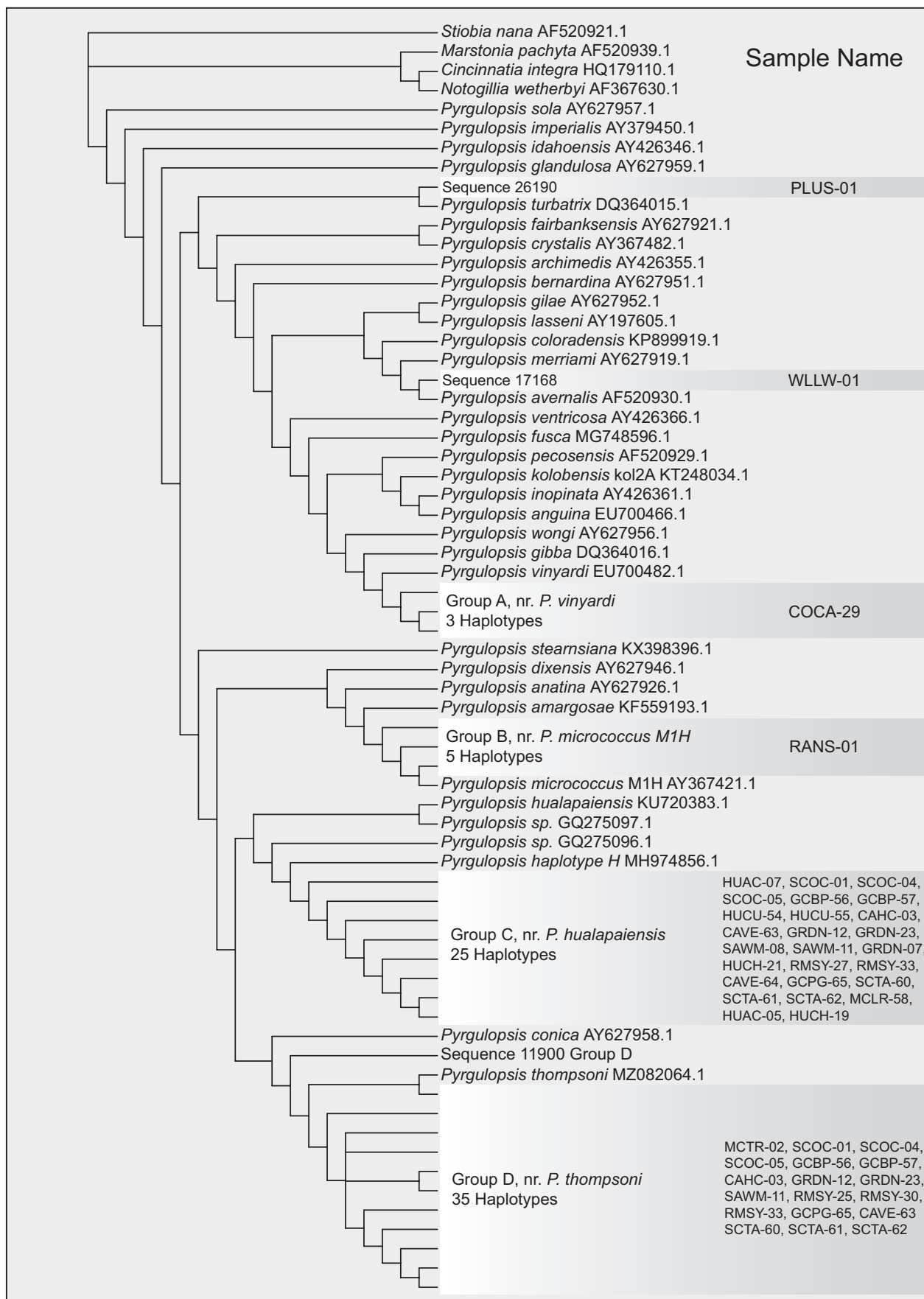


FIGURE 1 The maximum likelihood phylogeny estimated from 137 COI sequences obtained from GenBank and the 70 ASVs recovered in this study. Taxa were pared from the tree for visualization purposes. Most of the ASVs cluster into four clades (Groups A–D), identified by the nearest named species included in the phylogeny. Where multiple haplotypes were not identified for a particular ASV the sequence ID is given in bold (*P. aernalis* and *P. turatrix*).

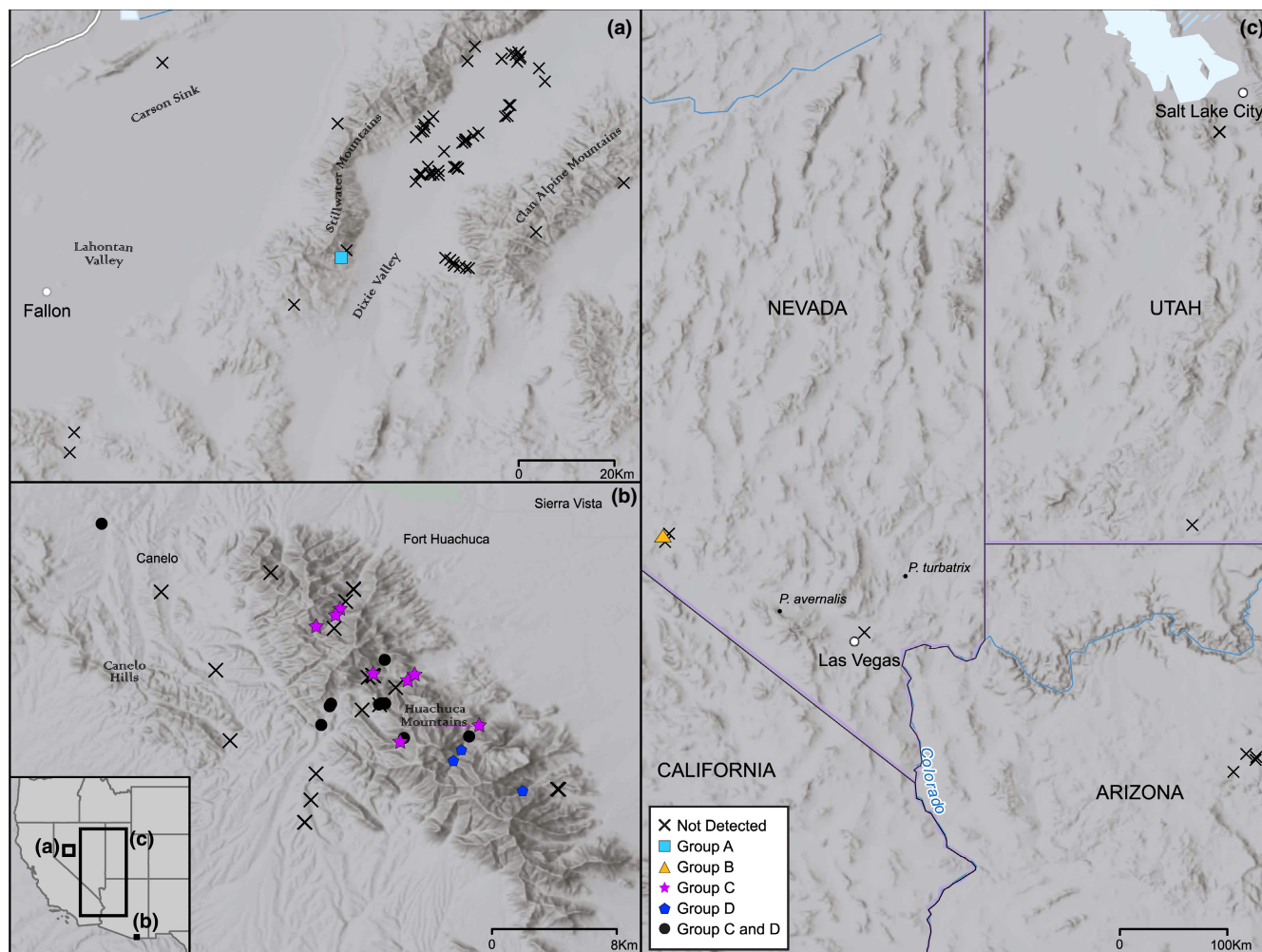


FIGURE 2 eDNA survey sites included in this study. Filled shapes represent sites where *Pyrgulopsis* was detected while “X” represents no detection. Panel ‘a’ depicts sampled sites around Fallon, NV area. Panel ‘b’ depicts sites sampled in the Huachuca, AZ region. Panel ‘c’ depicts sites across the Great Basin, in Utah, Nevada, and Arizona. Service Layer Credits: ESRI, HERE, Garmin, (c) OpenStreetMap contributors, and the GIS user community.

which four chimeric sequences and two reverse complemented sequences were excluded from downstream analyses. Inspection of the remaining 121 ASVs revealed a subset of 15 detections with low read abundance (<50 reads) that were subjected to further inspection.

Among these low read abundance ASVs, there were two cases where a distinct ASV was detected at only a single site. The first of these was a close match to *P. aernalis*, the Moapa Pebblesnail (99.65% identity, 39 reads, sequence I.D. 17168, Table S1). This sequence was found at Plummer Springs (PLUS-01) on the Moapa Valley National Wildlife Refuge and was retained as a true positive as it is known from this area (Hershler, 1998; Hershler & Liu, 2017; Ledbetter et al., 2023). The other ASV with low abundance (15

reads, sequence I.D. 26190, Table S1) was retained as a true positive because it had 100% sequence identity with *P. turatrix*, the Southwest Nevada Pyrg, and was collected from Willow Springs (WLLW-01), a site from which the species had been previously collected (Hershler, 1998; Ledbetter et al., 2023). For the remaining *Pyrgulopsis* ASVs with abundance <50 reads, detected at only a single site, the ASV differed by only 1–2 base pairs from a more common haplotype or could be traced to a background contamination event (see below). We conservatively elected to remove these ASVs ($n = 13$) as potential false positives.

Our BLAST search of 2955 COI sequences previously processed in the lab identified nine sites (six species, 38 ASVs) where an ASV shared 100% identity between sites that were separated by

>500 miles. Further inspection revealed these sites were processed in the same batch of amplifications, leading us to suspect these results were due to background contamination in a common PCR reagent. We elected to remove detections at these sites for the ASVs incongruent with known species distributions.

The remaining 70 ASVs primarily clustered into four clades (arbitrarily designated here as Groups A–D): Group A, the nr.-*P. vinyardi* clade ($n=3$ ASVs); Group B, the nr.-*P. micrococcus* M1H clade ($n=5$ ASVs); Group C, the nr.-*P. hualapaiensis* clade ($n=25$ ASVs); and Group D, the nr.-*P. thompsoni* clade ($n=35$ ASVs). These groups, further described using the names of their sister taxon in Figure 1, clustered together with each group containing ASVs that differed from another ASV or the sister taxon at only a few nucleotide sites (usually 1–2). In order to aid in visualization, these groups are collapsed in size in Figure 1, and represented by blocks labeled with locality information. Group A consists of three haplotypes found at Coyote Canyon Spring (39.545038, -118.225527) a small drainage on the edge of the Dixie Valley in central Nevada. These haplotypes are closely related to *P. vinyardi* (99.66%, isolate: SM38A, GenBank: [EU700482](#)) and *P. gibba* (99.66%, isolate: P134D, GenBank: [AY426359](#)) where *P. gibba* has been previously collected in Churchill county, near Dixie Valley (Springsnail Conservation Team, 2020), but to our knowledge this is the first detection of this species at this locality.

Group B is comprised of five sequences found at site 7J Ranch-9 Spring (RANS-01) identical or nearly identical to *P. micrococcus* M1H. 7J Ranch-9 Spring is located in southwestern Nevada (37.03795, -116.7123) near the area where *P. micrococcus* M1H was first discovered (Liu et al., 2003) (Figure 2). Group C haplotypes, first identified by Hurt (2004), were detected at 26 sites (See Table S1) all within the Huachuca region (See Figure 2 map). Many of these sites overlap with previous detections of this lineage from the same area (Hershler et al., 2016; Hurt, 2004). It is a similar case for Group D, comprised of 35 ASVs identical to or nearly identical to *P. thompsoni*, the Huachuca Springsnail. These 35 ASVs are distributed among 18 different sites throughout the Huachuca region, many of which overlap with sites where Group C also occurs (See Figure 2 and Table S2 for details). *Pyrgulopsis thompsoni* is described from this region (Hershler & Landye, 1988) and has been previously sampled from sites included in this study.

3.2 | Comparison with conventional sampling

Conventional sampling and eDNA sampling concurred on the presence of *Pyrgulopsis* at all but one site. *Pyrgulopsis* were not detected by either method at five sites. At 12 sites in which conventional sampling detected *Pyrgulopsis*, 24 of the 26 eDNA samples were positive (92.3% detection rate), including at least one replicate at every site (Table S3). At one site, eDNA sampling detected *Pyrgulopsis* where none had been collected with conventional sampling. This site, however, was in the same basin and <2km from a site with a previous conventional detection. Because this site, SC02 was weakly positive for *Pyrgulopsis* eDNA (four reads) and only 50m downstream

from site SC01, which was a strongly positive eDNA site; this detection may represent a degraded eDNA signal from the upstream population.

4 | DISCUSSION

Here we performed the first, to our knowledge, application of eDNA sampling for *Pyrgulopsis* springsnails. Environmental DNA sampling and sequencing detected springsnails of the genus *Pyrgulopsis* (and likely closely related genera) with results highly concordant with conventional sampling strategies. Among the 142 sites sampled here, 34 resulted in detections (24%), with the majority of these detections concordant with historical understandings of species distributions. For example, where springsnails were detected in the most recent netting-based survey in the Huachuca landscape, 94.4% were also positive with eDNA sampling. Additional detections with eDNA sampling on this landscape were consistent with eDNA transport from a nearby, hydrologically connected site with known historical species occurrence. Further, the performance of *Pyrgulopsis* eDNA sampling here is consistent with the literature on morphologically similar invasive hydrobiid New Zealand Mudsnaills (*Potamopyrgus antipodarum*) (Goldberg et al., 2013; Ponce et al., 2021; Woodell et al., 2021). Multiple studies have found that eDNA sampling for *P. antipodarum* can be effective, even for low density population detection. For example, Goldberg et al. (2013), using a taxon-specific qPCR assay, were successful in detection of *P. antipodarum* at densities as low as one individual per 1.5L of water. In natural systems, Ponce et al. (2021) reported concordance in eDNA sampling and conventional kick-net sampling for *P. antipodarum* across six streams in Washington (USA) and Woodell et al. (2021) reported detection of a new invasion with eDNA sampling in Pennsylvania (USA), which was later confirmed with conventional sampling.

Our detection results for low-density populations might be further improved by optimizing the field sampling design. Particularly in complex habitats, detection can be limited by sampling proximity to patchily distributed taxa because of limited spatial dispersion of the eDNA signal (Bedwell & Goldberg, 2020; Brandão-Dias et al., 2023; Shogren et al., 2017; Wilcox et al., 2015). When the location of optimal habitat for the focal taxon is unknown, detection might be accomplished by repeated sampling over space, but when the preferred habitat is known, then sampling might target those locations (Bedwell & Goldberg, 2020). *Pyrgulopsis* are often closely associated with spring heads, so while subsampling over space may be effective (Bedwell & Goldberg, 2020), it may also be effective to simply target spring heads in complex habitats or springs with large pool habitats. An additional consideration is sample timing relative to spring hydroperiod. Sampling for eDNA requires sufficient surface water, but sampling during wet periods could also result in wasted effort on ephemeral waters that do not represent suitable springsnail habitat.

When sampling was scaled up to encompass entire landscapes, we observed that good detection rates at the local scale, translated

into landscape patterns of detection that were largely concordant with previous conventional sampling. In the Huachuca region, we detected *Ptho5* (and associated haplotypes Group C, Figure 1) as well as *P. thompsoni* (and associated haplotypes Group D, Figure 1), both taxa previously collected from springs in this area (Hershler & Landye, 1988; Hurt, 2004). Additionally, our comparisons with sites for which conventional sampling data was available were 94.4% concordant (17/18 sites) with the only discrepancy coming from a weakly positive eDNA detection (four reads) at a location where springsnails had not been previously reported. High concordance (>90%) among sites that are known to host springsnails and sites that do not is an indicator that the assay is performing as intended. In the Fallon region, we detected snails from the *P. vinyardi/gibba* species complex. While we are not aware of a previous detection from this exact site, this finding is unsurprising in that *P. gibba* was previously collected at many sites in the area and is common throughout central Nevada and into California (Springsnail Conservation Team, 2020). The three remaining species detected were from other parts of southern Nevada outside of the Nellis AFB region. *Pyrgulopsis micrococcus* isolate M1H and associated haplotypes (Group B, Figure 1) were isolated from a spring (RANS-01) in a region where multiple *P. micrococcus* lineages have been identified (Liu et al., 2003). The remaining two detections were of low abundance (<50 reads) but from springs where these species have been previously reported. *Pyrgulopsis avernalis*, the Moapa Valley Springsnail, was detected in Plummer Spring (PLUS-01) in the Moapa Valley National Wildlife Refuge, the area for which the species is named. The final low abundance detection was of *P. turbatrix* from Willow Springs (WLLW-01), a site from which this species has been previously collected via conventional sampling and was thus unsurprising. While these ASVs with low abundance reads are likely true positives, these detections should be verified with additional sampling.

The uncertainty associated with some detections in this study is not unique. Environmental DNA metabarcoding for rare species currently faces a fundamental signal-to-noise problem: Both true presence of rare species and low-level contamination tend to result in low read abundances. Applying thresholds that discard reads below some threshold abundance is a common approach to remove contamination (Drake et al., 2022), but this practice can also result in the loss of true detections and impact ecological inference (Littleford-Colquhoun et al., 2022; Young et al., 2013). In this study, suspected contamination events tended to have low read abundances (mean=892 range=1–17,812 reads/library in those discarded), but this overlapped putative real detections (mean=9133, range=15–326,382 reads/library), including several with very low read counts that were almost certainly accurate. For example, we detected just 39 reads for a haplotype very similar to a reference for *P. avernalis*, the Moapa Pebblesnail (99.65% identity) at Plummer Springs on the Moapa Valley National Wildlife Refuge. This was almost certainly a true positive detection because (1) this taxon was never otherwise handled in the lab or detected in our study and (2) is known to occupy this area (Hershler, 1998; Hershler & Liu, 2017; Ledbetter et al., 2023). With current metabarcoding technologies and risks of

contamination during both library preparation and sequencing (i.e., tag switching and run-to-run carry-over), distinguishing true positive detections and contamination based on read abundance alone may not be possible. In this study, we took a conservative approach to inference of species presence by discarding low-abundance detections unless there was outside information suggesting species presence and discarding any other detections associated with a suspected library preparation contamination event. However, the ultimate solution for ambiguous or uncertain detections is likely to be repeat sampling. Persistent populations should result in persistent eDNA signals over time, even if those persistent signals are weak. Thus, repeated detection of low-abundance haplotypes over repeated sampling, library preparation, and sequencing efforts provides strong evidence of true species presence. Environmental DNA sampling may also play a complementary role to conventional sampling where rapid eDNA sampling can identify habitats for re-sampling with conventional approaches which could establish species presence with a higher level of certainty. Before making conservation management decisions based on specific detection data from this study, we recommend repeated sampling with eDNA or conventional approaches to confirm new detections.

Another uncertainty that is general to current eDNA metabarcoding tools concerns real versus error-derived haplotype diversity. A unique attribute of eDNA sampling is that the DNA of a potentially large number of individuals are present in a sample. In our system, our water samples could potentially contain DNA from hundreds of individuals, while most conventional sampling typically results in sequencing of just 1–30 individuals from a site (Liu et al., 2003; McKelvey et al., 2020). Thus, there is potential to detect greater haplotype diversity, including rare haplotypes. Simultaneously, it is well known that PCR and high-throughput sequencing generate errors which can be difficult to distinguish from real biological diversity. Historically, operational taxonomic unit (OTU) approaches have compressed both real biological and sequencing error-derived diversity by representing a range of haplotypes with centroid or representative sequences. In microbial systems, this approach resulted in the loss of meaningful biological inference at fine scales, motivating the development of denoising algorithms which attempt to correct for sequencing error to estimate original template sequences. However, even the best performing algorithms generate some spurious reads (i.e., inference of haplotypes with sequencing errors; Callahan et al., 2016). The diversity of haplotypes detected here likely represent a composite of both real biological diversity detected due to deep sampling of individuals from environmental samples and sequencing errors. In cases where a species was particularly abundant, we observe a primary haplotype that accounts for the majority of the total reads recovered (~85%–95%), with one or two minor haplotypes accounting for most of the remaining reads followed by many rare haplotypes of low abundance (see Table S1, Groups C and D). These rare haplotypes often differ by only 1–2 base pairs from the primary haplotype with the mutation often being a non-synonymous change. Since the majority of differences we

observed in our alignment of 131 *Pyrgulopsis* species are synonymous substitutions, it stands to reason that these rare haplotypes with non-synonymous changes could represent sequencing errors. However, these rare haplotypes may also represent somatic mutations within large populations that are not typically observed using Sanger sequencing.

With an appreciation of the potential limitations, this study suggests several ways in which eDNA sampling can be a useful tool in the toolbox for research and monitoring of *Pyrgulopsis* and other cryptic, species-rich, and incompletely described taxa. Compared to conventional sampling, eDNA sampling is more time efficient, requires minimal personnel training, and is likely more cost effective for species detection. We have found that trained field technicians can struggle to distinguish even deeply divergent mollusk taxa (e.g., *Pyrgulopsis* versus immature *Physa*); >400 MYA divergence (McKelvey et al., 2020). For these cryptic taxa, molecular analyses are still required. A typical conventional sampling approach for *Pyrgulopsis* may require a field crew to net 10–30 individuals, then submit these for barcoding. At a cost of \$30/barcode, this translates into a cost of at least \$300/site. Using a MiSeq platform, we estimate our cost/sample in this study to be approximately \$150. With increased sample multiplexing, samples could be sequenced on a higher capacity platform, further reducing the per sample cost. Furthermore, existing eDNA sampling protocols (e.g., Carim, Dysthe, et al., 2016; Carim, McKelvey, et al., 2016) can be incorporated into conventional springsnail surveys or piggy-backed with non-springsnail surveys in remote areas, reducing the logistical cost of sampling.

With the increasing use of eDNA methods over the last two decades has come an increasingly complete suite of tools and publications addressing eDNA sampling best practices, project planning, and application (e.g., Goldberg et al., 2016; Kelly et al., 2023; Liu et al., 2020; Maloy et al., 2023) to help guide potential users seeking to partner with molecular labs.

This study demonstrates that partnerships between existing springs inventory and molecular laboratories can deploy eDNA sampling as a rapid, low-cost way to identify putative occupied sites (and rare species or haplotypes) for further examination. Just as conventional barcoding has revolutionized our view of cryptic biodiversity at global scales (Hajibabaei et al., 2007), we anticipate that eDNA sampling will play an increasingly important role in characterizing unseen diversity at landscape scales.

AUTHOR CONTRIBUTIONS

Daniel D. Vanderpool: Conceptualization (equal); data curation (lead); formal analysis (lead); investigation (lead); methodology (equal); supervision (supporting); writing – original draft (lead); writing – review and editing (equal). **Taylor M. Wilcox:** Conceptualization (equal); formal analysis (supporting); funding acquisition (lead); investigation (equal); methodology (equal); project administration (lead); supervision (lead); writing – original draft (equal); writing – review and editing (equal). **Michael K. Young:** Methodology (supporting); writing – review and editing (equal). **Kristine L. Pilgrim:** Data

curation (supporting); methodology (supporting); resources (supporting); supervision (supporting); validation (supporting). **Michael K. Schwartz:** Funding acquisition (supporting); project administration (supporting); resources (supporting); supervision (supporting); writing – review and editing (supporting).

ACKNOWLEDGMENTS

The sampling for this study was made possible by contributions of many different field partners including Jeff Sorensen (Arizona Game and Fish), Rick Lance (Army Corps of Engineers Environmental Lab), Rodrigo Seden, Alison Ochoa, Joanna Elmore, Jennifer Hernandez (USDA Forest Service), Andrea Hazelton, and Larry Stevens (Springs Stewardship Institute). This project was only possible due to the collaboration of Department of Defense partners who provided access and sampling permits, historic survey information, and in some cases collected eDNA samples or accompanied sampling crews including Debbie Brewer (Fort Huachuca), Anna Johnson (Nellis Air Force Base), Ann Schofield (Fallon Naval Air Station), Chris Frauenhofer (Camp Williams), and Hannah Telle (Camp Navajo). We also benefited from logistical coordination provided by John Kraft (Coronado National Forest), Rhonda Stewart, Larry Cardova (Lincoln National Forest), and access to Ramsey Canyon Preserve facilitated by Peter Lieterman (The Nature Conservancy). Primer testing benefited from collection loans from the American Museum of Natural History (Catalog Numbers 883262, 905339, 873231, 1135056, 883553, and 1152489) as well as loans from Kathryn Perez and Texas Parks & Wildlife Department. This work was funded by the Strategic Environmental Resource Development Program (SERDP) Project RC18-1348 and by the USDA Forest Service Rocky Mountain Research Station. Sequencing was performed at the University of Montana Genomics Core Facility supported in part by the NIGMS CoBRE Grant P20GM103546.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

All FASTQ files associated with this work are archived under SRA BioProject ID: PRJNA1020154. Supplementary Material is available from Dryad: <https://doi.org/10.5061/dryad.tb2rbp07q>.

ORCID

Daniel D. Vanderpool  <https://orcid.org/0000-0002-6856-5636>

Taylor M. Wilcox  <https://orcid.org/0000-0003-3341-7374>

Michael K. Young  <https://orcid.org/0000-0002-0191-6112>

REFERENCES

- Bedwell, M. E., & Goldberg, C. S. (2020). Spatial and temporal patterns of environmental DNA detection to inform sampling protocols in lentic and lotic systems. *Ecology and Evolution*, 10(3), 1602–1612.
- Brandão-Dias, P. F. P., Hallack, D. M. C., Snyder, E. D., Tank, J. L., Bolster, D., Volponi, S., Shogren, A. J., Lamberti, G. A., Bibby, K., & Egan, S. P. (2023). Particle size influences decay rates of environmental DNA in aquatic systems. *Molecular Ecology Resources*, 23(4), 756–770.

- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13(7), 581–583.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, 10(1), 421.
- Carim, K. J., Dysthe, J. C. S., Young, M. K., McKelvey, K. S., & Schwartz, M. K. (2016). An environmental DNA assay for detecting Arctic grayling in the upper Missouri River basin, North America. *Conservation Genetics Resources*, 8(3), 197–199.
- Carim, K. J., McKelvey, K. S., Young, M. K., Wilcox, T. M., & Schwartz, M. K. (2016). A protocol for collecting environmental DNA samples from streams. RMRS-GTR-355. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.
- Chernomor, O., von Haeseler, A., & Minh, B. Q. (2016). Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biology*, 65(6), 997–1008.
- Clusa, L., Ardura, A., Gower, F., Miralles, L., Tsartsianidou, V., Zaiko, A., & Garcia-Vazquez, E. (2016). An easy phylogenetically informative method to trace the globally invasive *Potamopyrgus* mud snail from river's eDNA. *PLoS One*, 11(10), e0162899.
- Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D. M., de Vere, N., Pfrender, M. E., & Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, 26(21), 5872–5895.
- Drake, L. E., Cuff, J. P., Young, R. E., Marchbank, A., Chadwick, E. A., & Symondson, W. O. C. (2022). An assessment of minimum sequence copy thresholds for identifying and reducing the prevalence of artefacts in dietary metabarcoding data. *Methods in Ecology and Evolution*, 13(3), 694–710.
- Elbrecht, V., & Leese, F. (2017). Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. *Frontiers of Environmental Science & Engineering in China*, 5, 11. <https://doi.org/10.3389/fenvs.2017.00011>
- Eloe-Fadrosh, E. A., Ivanova, N. N., Woyke, T., & Kyrpides, N. C. (2016). Metagenomics uncovers gaps in amplicon-based detection of microbial diversity. *Nature Microbiology*, 1(4), 15032.
- Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3(5), 294–299.
- Geller, J., Meyer, C., Parker, M., & Hawk, H. (2013). Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Molecular Ecology Resources*, 13(5), 851–861.
- Goldberg, C. S., Turner, C. R., & Deiner, K. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution*, 7, 1299–1307.
- Goldberg, C. S., Sepulveda, A., Ray, A., Baumgardt, J., & Waits, L. P. (2013). Environmental DNA as a new method for early detection of New Zealand Mudsnailed (*Potamopyrgus antipodarum*). *Freshwater Science*, 32(3), 792–800.
- Hajibabaei, M., Singer, G. A. C., Hebert, P. D. N., & Hickey, D. A. (2007). DNA barcoding: How it complements taxonomy, molecular Phylogenetics and population genetics. *Trends in Genetics*, 23(4), 167–172.
- Hershler, R. (1998). A systematic review of the Hydrobiid snails (Gastropoda: Rissoidae) of the Great Basin, Western United States. Part 1. Genus *Pyrgulopsis*. *The Veliger*, 41, 1–132.
- Hershler, R., & Landye, J. J. (1988). Arizona Hydrobiidae (Prosobranchia: Rissoidae). *Smithsonian Contributions to Zoology*, 459, 1–63.
- Hershler, R., & Liu, H.-P. (2017). *Annotated checklist of freshwater truncatelloidean gastropods of the Western United States, with an illustrated key to the genera*. U.S. Department of Interior, Bureau of Land Management.
- Hershler, R., Liu, H.-P., & Howard, J. (2014). Springsnails: A new conservation focus in Western North America. *Bioscience*, 64(8), 693–700.
- Hershler, R., Liu, H.-P., & Stevens, L. E. (2016). A new Springsnail (Hydrobiidae: *Pyrgulopsis*) from the lower Colorado River basin, northwestern Arizona. *Western North American Naturalist/Brigham Young University*, 76(1), 72–81.
- Hershler, R., Ratcliffe, V., Liu, H.-P., Lang, B., & Hay, C. (2014). Taxonomic revision of the *Pyrgulopsis gilae* (Caenogastropoda, Hydrobiidae) species complex, with descriptions of two new species from the Gila River basin, New Mexico. *ZooKeys*, 429, 69–85.
- Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, 35(2), 518–522.
- Hurt, C. R. (2004). Genetic divergence, population structure and historical demography of rare Springsnails (*Pyrgulopsis*) in the lower Colorado River basin. *Molecular Ecology*, 13(5), 1173–1187.
- Johnson, P. D., Bogan, A. E., Brown, K. M., Burkhead, N. M., Cordeiro, J. R., Garner, J. T., Hartfield, P. D., Lepitzki, D. A. W., Mackie, G. L., Pip, E., Tarpley, T. A., Tiemann, J. S., Whelan, N. V., & Strong, E. E. (2013). Conservation status of freshwater gastropods of Canada and the United States. *Fisheries*, 38(6), 247–282.
- Kalyanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermini, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14(6), 587–589.
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772–780.
- Kelly, R. P., Lodge, D. M., Lee, K. N., Theroux, S., Sepulveda, A. J., Scholin, C. A., Craine, J. M., Allan, E. A., Nichols, K. M., Parsons, K. M., Goodwin, K. D., Gold, Z., Chavez, F. P., Noble, R. T., Abbott, C. L., Baerwald, M. R., Naum, A. M., Thielen, P. M., Simons, A. L., ... Weisberg, S. B. (2023). Toward a national eDNA strategy for the United States. *Environmental DNA*, 6(1), e432. <https://doi.org/10.1002/edn3.432>
- Kelly, R. P., Shelton, A. O., & Gallego, R. (2019). Understanding PCR processes to draw meaningful conclusions from environmental DNA studies. *Scientific Reports*, 9(1), 12133.
- Kronenberger, J. A., Wilcox, T. M., Mason, D. H., Franklin, T. W., McKelvey, K. S., Young, M. K., & Schwartz, M. K. (2022). eDNAAssay: A machine learning tool that accurately predicts qPCR cross-amplification. *Molecular Ecology Resources*, 22(8), 2994–3005.
- Lanfear, R., Calcott, B., Ho, S. Y. W., & Guindon, S. (2012). Partitionfinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, 29(6), 1695–1701.
- Lavretsky, P., Duenez, E., & Sorensen, J. (2021). Population genetics and conservation of recently discovered Springsnails in Arizona. *The Journal of Molluscan Studies*, 87(3), eyab020.
- Ledbetter, J. D., Stevens, L. E., Springer, A., & Brandt, B. (2023). *Springs online: A springs and springs-dependent taxa database*. Online Database Vers. 23.0. Springs Stewardship Institute.
- Littleford-Colquhoun, B. L., Freeman, P. T., Sackett, V. I., Tulloss, C. V., McGarvey, L. M., Geremia, C., & Kartzinel, T. R. (2022). The precautionary principle and dietary DNA metabarcoding: Commonly used abundance thresholds change ecological interpretation. *Molecular Ecology*, 31(6), 1615–1626.
- Liu, H.-P., Hershler, R., & Clift, K. (2003). Mitochondrial DNA sequences reveal extensive cryptic diversity within a Western American Springsnail. *Molecular Ecology*, 12(10), 2771–2782.
- Liu, H.-P., Hershler, R., & Hovingh, P. (2017). Molecular evidence enables further resolution of the Western North American *Pyrgulopsis kolobensis* complex (Caenogastropoda: Hydrobiidae). *The Journal of Molluscan Studies*, 84(1), 103–107.
- Liu, M., Clarke, L. J., Baker, S. C., Jordan, G. J., & Burridge, C. P. (2020). A practical guide to DNA metabarcoding for entomological ecologists. *Ecological Entomology*, 45, 373–385.

- Maloy, A., Rees, C., Bockrath, K., Vazquez, J., Russ, O., Grote, A., Mize, E., Tuttle-Lau, M., Bloomfield, J., Olsen, J., & Benson, A. (2023). *Environmental DNA (eDNA) best management practices for project planning, deployment, and application*. U.S. Fish and Wildlife Service. https://www.researchgate.net/publication/373894943_Environmental_DNA_eDNA_Best_Management_Practices_for_Project_Planning_Deployment_and_Application
- McKee, A. M., Spear, S. F., & Pierson, T. W. (2015). The effect of dilution and the use of a post-extraction nucleic acid purification column on the accuracy, precision, and inhibition of environmental DNA samples. *Biological Conservation*, 183, 70–76.
- McKelvey, K. S., Kallstrom, C., Ledbetter, J., Sada, D. W., Pilgrim, K. L., & Schwartz, M. K. (2020). An inventory of Springsnails (*Pyrgulopsis* spp.) in and adjacent to the Spring Mountains, Nevada. *Western North American Naturalist*, 80(2), 183–193.
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, 37(5), 1530–1534.
- Ponce, J. J., Arismendi, I., & Thomas, A. (2021). Using in-situ environmental DNA sampling to detect the invasive New Zealand mud snail (*Potamopyrgus antipodarum*) in freshwaters. *PeerJ*, 9, e11835.
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org>
- Shogren, A. J., Tank, J. L., Andruszkiewicz, E., Olds, B., Mahon, A. R., Jerde, C. L., & Bolster, D. (2017). Controls on eDNA movement in streams: Transport, retention, and resuspension. *Scientific Reports*, 7(1), 5065.
- Springsnail Conservation Team. (2020). *Conservation strategy for Springsnails in Nevada and Utah, version 1.0*. Springs Stewardship Institute.
- Stevens, L. E., Holcomb, K., Crookshanks, C., Sada, D. W., Jenness, J., & Szabo, K. (2022). A strategy for conservation of Springsnails in Nevada and Utah, USA. *Sustainability: Science, Practice and Policy*, 14(15), 9546.
- Tullo, E. M., & Sorensen, J. A. (2021). *Five-Year Progress Report on Huachuca Springsnail Candidate Conservation Agreement Actions, 2016–2020, Nongame And Endangered Wildlife Program Technical Report 343*. Arizona Game and Fish Department, Phoenix, Arizona.
- Vanderpool, D., Wilcox, T. M., Young, M. K., Pilgrim, K. L., & Schwartz, M. K. (2023a). Supplementary data for “simultaneous species detection and discovery with environmental DNA Metabarcoding: A freshwater mollusk case study”. [dataset]. *Dryad*. <https://doi.org/10.5061/dryad.tb2rbp07q>
- Vanderpool, D., Wilcox, T. M., Young, M. K., Pilgrim, K. L., & Schwartz, M. K. (2023b). COI metabarcoding reads for *Pyrgulopsis* Springsnails. Sequence read archive; BioProject ID: PRJNA1020154. [dataset].
- Wilcox, T. M., McKelvey, K. S., Young, M. K., Sepulveda, A. J., Shepard, B. B., Jane, S. F., Whiteley, A. R., Lowe, W. H., & Schwartz, M. K. (2016). Understanding environmental DNA detection probabilities: A case study using a stream-dwelling char *Salvelinus fontinalis*. *Biological Conservation*, 194, 209–216.
- Wilcox, T. M., McKelvey, K. S., Young, M. K., Lowe, W. H., & Schwartz, M. K. (2015). Environmental DNA particle size distribution from brook trout (*Salvelinus fontinalis*). *Conservation Genetics Resources*, 7(3), 639–641.
- Woodell, J. D., Neiman, M., & Levri, E. P. (2021). Matching a snail's pace: Successful use of environmental DNA techniques to detect early stages of invasion by the destructive New Zealand mud snail. *Biological Invasions*, 23(10), 3263–3274.
- Young, M. K., McKelvey, K. S., Pilgrim, K. L., & Schwartz, M. K. (2013). DNA barcoding at riverscape scales: Assessing biodiversity among fishes of the genus *Cottus* (Teleostei) in northern Rocky Mountain streams. *Molecular Ecology Resources*, 13(4), 583–595.
- Young, M. K., Smith, R., Pilgrim, K. L., & Schwartz, M. K. (2021). Molecular species delimitation refines the taxonomy of native and nonnative physine snails in North America. *Scientific Reports*, 11(1), 21739.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Vanderpool, D. D., Wilcox, T. M., Young, M. K., Pilgrim, K. L., & Schwartz, M. K. (2024). Simultaneous species detection and discovery with environmental DNA metabarcoding: A freshwater mollusk case study. *Ecology and Evolution*, 14, e11020. <https://doi.org/10.1002/ece3.11020>