Environmental DNA metabarcoding of fish in the Central Arizona Project canal during water release from Lake Pleasant

Gila River Basin Native Fishes Conservation Program
Lower Colorado Region
EcoLab-LCUAS-2024-03

U.S. Department of the Interior **February 2024**

Mission Statements

The U.S. Department of the Interior protects and manages the Nation's natural resources and cultural heritage; provides scientific and other information about those resources; honors its trust responsibilities or special commitments to American Indians, Alaska Natives, and affiliated Island Communities.

The mission of the Bureau of Reclamation is to manage, develop, and protect water and related resources in an environmentally and economically sound manner in the interest of the American public.

Disclaimer – The information provided in this report is believed to be appropriate and accurate for the specific purposes described herein, but users bear all responsibility for exercising sound engineering judgment in its application, especially to situations different from those studied. References to commercial products do not imply endorsement by the Bureau of Reclamation and may not be used for advertising or promotional purposes.

Cover Photo – Lake Pleasant, AZ. The photo is cropped from an image publicly available from Wikimedia Commons (https://commons.wikimedia.org/wiki/File:Lake_Pleasant,_Arizona_03.jpg) and is used under a Creative Commons 3.0 license (https://creativecommons.org/licenses/by-sa/3.0/deed.en).

Environmental DNA metabarcoding of fish in the Central Arizona Project canal during water release from Lake Pleasant

Gila River Basin Native Fishes Conservation Program
Lower Colorado Region
EcoLab-LCUAS-2024-03

Prepared by:

Yale Passamaneck, Ph.D.
Bureau of Reclamation
Technical Service Center
Denver, Colorado

U.S. Department of the Interior **February 2024**

Environmental DNA metabarcoding of fish in the Central Arizona Project canal during water release from Lake Pleasant

Gila River Basin Native Fishes Conservation Program
Lower Colorado Region
EcoLab-LCUAS-2024-03

Prepared by: Yale Passamaneck, Ph.D., Ecological Research Laboratory, Hydraulic Investigations and Laboratory Services, Technical Service Center, Bureau of Reclamation

Peer Review Certification

This section has been reviewed and is believed to be in accordance with the service agreement and standards of the profession.

Peer reviewed by: Jacque Keele, Ph.D., Ecological Research Laboratory, Hydraulic Investigations and Laboratory Services, Technical Service Center, Bureau of Reclamation

Acronyms and Abbreviations

ASV	amplified sequence variant
CAP	Central Arizona Project
DNA	deoxyribonucleic acid
ESA	Endangered Species Act
eDNA	environmental DNA
PCR	polymerase chain reaction
Reclamation	Bureau of Reclamation
USFWS	U.S. Fish and Wildlife Service

Contents

Page

I. Introduction 1
II. Methods 3
  A. Sample collection 3
  B. PCR amplification 4
  C. Negative controls 5
  D. Sequencing 5
  E. Analysis 5
    1. DNA sequence data processing 5
  F. Taxonomic assignment 5
III. Results and Discussion 6
  A. CAP and Lake Pleasant eDNA sampling – Summer 2022 6
    1. Grass carp (*Ctenopharyngodon idella*) and common carp (*Cyprinus carpio*) 6
    2. Striped bass (*Morone saxatilis*) 8
    3. Channel catfish (*Ictalurus punctatus*) and Blue catfish (*Ictalurus furcatus*)...9
    4. Green sunfish (*Lepomis cyanellus*), Bluegill (*Lepomis macrochirus*), and Redear sunfish (*Lepomis microlophus*) 11
    5. Inland silverside (*Menidia beryllina*) 14
    6. Smallmouth bass (*Micropterus dolomieu*) and Largemouth bass (*Micropterus salmoides*) 15
    7. Gizzard shad (*Dorosoma cepedianum*) and Threadfin shad (*Dorosoma petenense*) 17
    8. Western mosquitofish (*Gambusia affinis*) 19
    9. Tilapia (*Oreochromis* sp.) 20
    10. Flathead catfish (*Pylodictis olivaris*) 21
    11. Salmon/trout (*Oncorhynchus* spp.) 22
    12. Non-fish sequences 23
  B. Distributions in Lake Pleasant, Waddell Canal and the CAP Canal 24
  C. Seasonal differences in eDNA 25
    1. Upstream of Waddell Canal and Lake Pleasant 25
    2. Downstream of Waddell Canal and Lake Pleasant 28
IV. References 31

Figures

Figure Page

Figure 1: Locations of sampling sites along the CAP Canal Waddell Canal, and Lake Pleasant...4
Figure 2: Distribution of grass carp (*C. idella*) sequence detections. 7
Figure 3: Distribution of common carp (*C. carpio*) sequence detections. 8
Figure 4: Distribution of striped bass (*M. saxatilis*) sequence detections. 9
Figure 5: Distribution of channel catfish (*I. punctatus*) sequence detections. 10
Figure 6: Distribution of blue catfish (*I. furcatus*) sequence detections. 11
Figure 7: Distribution of green sunfish (*L. cyanellus*) sequence detections. 12
Figure 8: Distribution of bluegill (*L. macrochirus*) sequence detections. 13
Figure 9: Distribution of redear sunfish (*L. microlophus*) sequence detections. 14
Figure 10: Distribution of inland silverside (*M. beryllina*) sequence detections. 15
Figure 11: Distribution of smallmouth bass (*M. dolomieu*) sequence detections. 16
Figure 12: Distribution of largemouth bass (*M. salmoides*) sequence detections. 17
Figure 13: Distribution of gizzard shad (*D. cepedianum*) sequence detections. 18
Figure 14: Distribution of threadfin shad (*D. petenense*) sequence detections. 19
Figure 15: Distribution of Western mosquitofish (*G. affinis*) sequence detections. 20
Figure 16: Distribution of tilapia (*Oreochromis* sp.) sequence detection. 21
Figure 17: Distribution of flathead catfish (*P. olivaris*) sequence detections. 22
Figure 18: Distribution of salmon/trout (*Oncorhynchus* spp.) sequence detection. 23
Figure 19: Percent read frequencies for fish species detected from eDNA metabarcoding from LAKEPLEA_01, WADD_PP, and WADD_01. 25
Figure 20: Percent read frequencies for fish species detected from eDNA metabarcoding from HASS_US. 26
Figure 21: Percent read frequencies for fish species detected from eDNA metabarcoding from HASS_PP. 26
Figure 22: Percent read frequencies for fish species detected from eDNA metabarcoding from HASS_DS. 27
Figure 23: Percent read frequencies for fish species detected from eDNA metabarcoding from HAWA_06. 27
Figure 24: Percent read frequencies for fish species detected from eDNA metabarcoding from WASG_01. 28
Figure 25: Percent read frequencies for fish species detected from eDNA metabarcoding from SALTGILA_US. 29
Figure 26: Percent read frequencies for fish species detected from eDNA metabarcoding from SALTGILA_PP. 29
Figure 27: Percent read frequencies for fish species detected from eDNA metabarcoding from SALTGILA_DS. 30

Appendices

Appendix 1: Table of sampling site information for eDNA metabarcoding survey……………......33

I. Introduction

The Central Arizona Project (CAP) canal is a 336-mile aqueduct which carries water from the Colorado River to central and southern Arizona. The CAP includes 14 pumping plants, which lift the water over 2,900 feet from the inlet in Lake Havasu to the terminus of the system near Tucson, Arizona. The system also includes a large storage reservoir (Lake Pleasant) which is operated by a hydroelectric pump/generating plant at New Waddell Dam. Typically, CAP water is pumped into Lake Pleasant during fall and winter, whereas water is released from the reservoir during spring and summer. The Bureau of Reclamation ('Reclamation' hereafter) initiated construction of the CAP in 1973, with water deliveries beginning in 1985 and construction being substantially completed in 1993.

Under Section 7 of the Endangered Species Act (ESA), Reclamation entered into formal consultation with the U.S. Fish and Wildlife Service (USFWS) over the potential for CAP water operations to impact federally listed species. Given that the CAP transports water between sub-basins of the Colorado River (from the Lower Colorado River basin to Gila River basin), concerns were raised regarding the potential of the CAP to transport non-native fishes between sub-basins which could in-turn travel upstream into waters inhabited by threatened and endangered native fishes. In 1994, USFWS issued a Biological Opinion (USFWS, 1994) and determined that the CAP jeopardized the existence of spikedace (*Meda fulgida*), loach minnow (*Tiaroga cobitis*), Gila topminnow (*Poeciliopsis occidentalis*), and razorback sucker (*Xyrauchen texanus*), and could adversely modify designated critical habitat of spikedace, loach minnow, and razorback sucker. Later revisions in 2001 and 2008, added Gila chub (*Gila intermedia*) and Chiricahua leopard frog (*Lithobates chiricahuensis*) as additional listed species affected by CAP operations (USFWS 2001, USFWS 2008).

In the 1994 Biological Opinion, the USFWS identified several reasonable and prude alternatives (RPAs) to remove jeopardy to these species – Reclamation later adopted these RPAs as Conservation Measures in the 2001 and 2008 revised Biological Opinions (USFWS, 2001; USWFS, 2008). One of the RPAs required Reclamation to develop and implement a long-term monitoring program to assess the presence and distribution of non-native fish in the CAP and its primary connected waters (canals and major streams) throughout the expected 100-year life of the CAP.

The long-term monitoring of the CAP and its primary connect waters was initiated in 1995, although pre-Opinion monitoring of the CAP occurred as early as 1986 (Mueller, 1996). Monitoring was conducted annually from 1995 through 2010; however, in recent years emphasis shifted towards monitoring wild populations of listed fishes in the Gila River basin. The CAP and its primary connected waters are now monitored once every 5 years according to Clarkson et al. (2011).

A previous study using environmental DNA metabarcoding was conducted in 2020 and 2021 to monitor for fish species at 83 sites along the length of the main CAP canal (Passamaneck, 2022).

In that study we detected sequences matching to 25 species of fish, although 6 of the identified species were attributed to DNA from frozen bait or were exotic and marine species whose origin in the samples could not be attributed to a specific live source. Direct comparisons between eDNA analysis and traditional sampling methods found similar trends between the eDNA read frequency for a species and the number of individuals caught by traditional sample. However, eDNA metabarcoding consistently detected more species at a given site than were caught by traditional methods.

The goal of the current study was to investigate the contribution of another input to the CAP canal system, the reservoir Lake Pleasant. Lake Pleasant has a capacity of over 800,000 acre feet, and is connected to the main CAP canal via the Waddell Canal. Water from the CAP canal is pumped into Lake Pleasant for storage during the fall and winter, and water is released back from Lake Pleasant to the CAP canal during the spring and summer to meet increased demand. At least 20 fish species have been reported as residing in Lake Pleasant (Stewart et al., 2007; Gill and Jones, 2019). Among these, white crappie (*Pomoxis annularis*), golden shiner (*Notemigonus crysoleucas*), and tilapia (*Orechromis* spp.) have been reported from Lake Pleasant but have not been caught in the CAP canal.

The current project collected samples from Lake Pleasant and the Waddell Canal, as well as from sites upstream and downstream of where the Waddell Canal meets the CAP canal. Sampling was conducted while water was being released from Lake Pleasant, in late August and early September. Sampling and analysis were conducted using the protocols as were in employed in the previous study of CAP canal fish eDNA, allowing for comparison of samples collected at the same sites during different flow regimes during winter and summer (Passamaneck, 2022).

II. Methods

A. Sample collection

Sample collection was based on the U.S. Forest Service protocol for eDNA collection from streams (Carim et al., 2016). Samples were collected by filtering water through Whatman glass microfiber filters, grade 934-AH, with a nominal particle retention size of 1.5 microns. Filters were placed in single use analytical filter funnels. Prior to filed collections, filters were individually packaged in sampling kits, along with nitrile gloves, sterile disposable forceps, and plastic baggies containing desiccant beads, for sample handling and storage. In the field the filter assemblies were attached to flexible hosing and a battery-powered peristaltic pump. At each sample site, the filter assembly was submerged in the sampled water and the pump was run until the targeted volume of filtrate (generally 2 liters) was collected in an outflow bucket. Following filtration, the filter assembly was recovered, and the filter was removed using gloved hands and sterile forceps. Each filter was placed in a desiccant baggie for preservation during storage and shipment. At each sampling site a field blank was collected, with one liter of distilled water filtered through the filter assembly, before the field samples were collected at the site. Three field samples were collected at each site. At pumping plants the samples were collected from three separate locations: the top (at the escape ladder upstream in the canal closest to the pumping plant; approximately 100 to 300 meters upstream of the pumping plant intakes depending on the site), bottom-right (at the escape ladder river-right in the forebay; approximately 10 meters from the pumping plant intakes), and bottom-left (at the escape ladder river-left in the forebay; approximately 10 meters from the pumping plant intakes) of the forebay.

Samples were collected between August 31, 2022, and September 2, 2022 (Appendix A). Sampling was conducted at 12 sites, including four sites in the CAP Canal upstream of Waddell Canal (HASS_US, HASS_PP [Hassayampa Pumping Plant], HASS_DS, and HAWA_06), four sites in the CAP Canal downstream of Waddell Canal (WASG_02, SALTGILA_US, SALTGILA_PP [Salt Gila Pumping Plant], and SALTGILA_DS), one site at the junction of the CAP Canal and the Waddell Canal (WASG_01), two sites in the Waddell Canal (WADD_PP [Waddell Pumping Plant] and WADD_01), and one site in Lake Pleasant (LAKEPLEA_01) (Figure 1; Appendix A).

Filters were processed for eDNA extraction in Reclamation's Ecological Research Laboratory (EcoLab) in Denver, CO. For each sample, half of the filter was processed for DNA extraction and purification, and the other half of the filter was stored at -80° C for subsequent analysis. All DNA extractions were performed using the Qiagen DNAeasy Blood & Tissue Kit. The proteinase K lysis was performed in Qiagen Investigator Lyse & Spin columns. Following an overnight incubation at 55° C, the lysate was recovered by centrifugation prior to further processing with the DNAeasy Blood & Tissue Kit. Following DNA extraction, the samples were purified using Zymo OneStep PCR Inhibitor Removal columns.

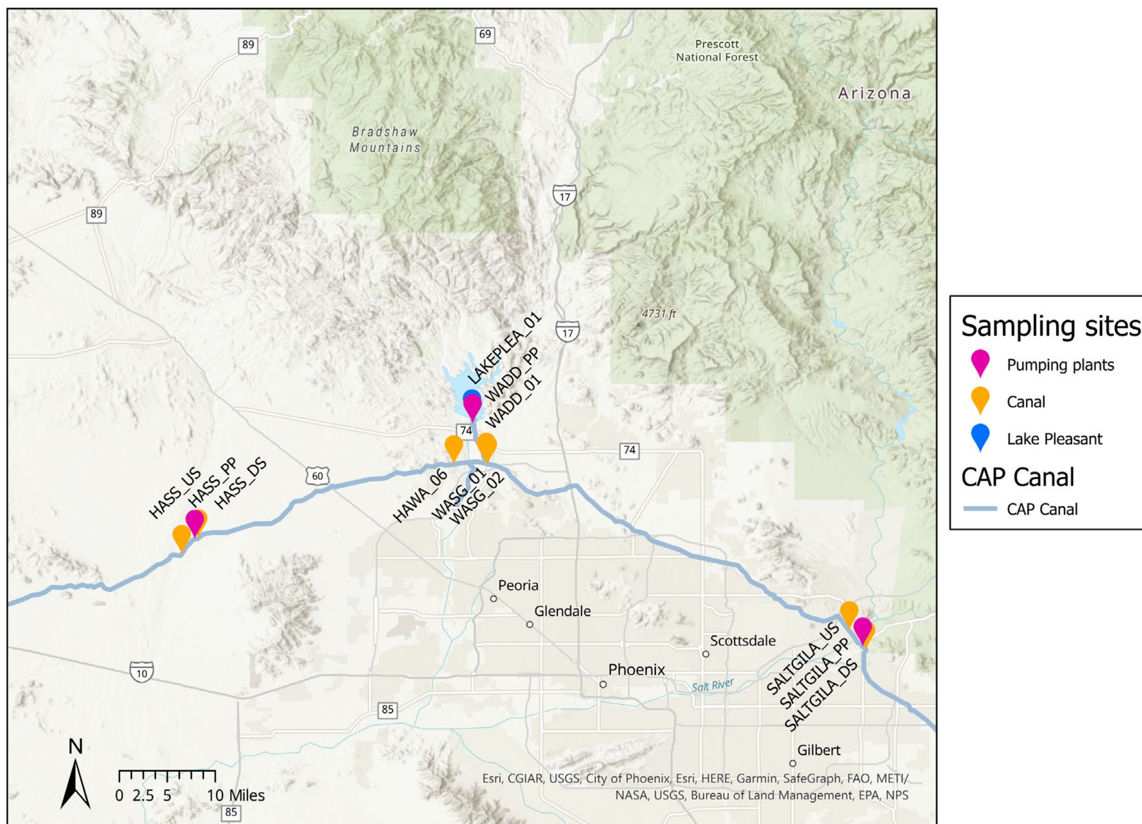

Figure 1: Locations of sampling sites along the CAP Canal, Waddell Canal, and Lake Pleasant.

B. PCR amplification

Polymerase chain reaction (PCR) amplification of DNA fragments was performed using the MiFish-U primers (Miya et al., 2015) which amplify a fragment of the 12S rRNA mitochondrial gene which is approximately 180 base pairs (bp) in length. For all samples, a first round PCR was performed using the MiFish-U primers. For samples intended for Illumina HiSeq sequencing, a second round of PCR was performed with MiFish-U primers labeled with unique 10 pb index sequences at the 5' end of the primer, to facilitate demultiplexing of DNA sequencing data. For samples intended for Amplicon-EZ sequencing, only the first round of PCR amplification with MiFish-U primers was performed. For all samples, PCR amplification was performed in four replicate reactions, with the replicates pooled prior to DNA sequencing. PCR reaction quality was checked by agarose gel electrophoresis. Reactions that did not show amplification were repeated to ensure the four replicates per sample were obtained. In some cases, sample dilution was adjusted to achieve amplification. All PCR amplifications were performed using Platinum SuperFi II Green MasterMix (Life Technologies). Following PCR amplification and agarose gel validation, PCR products were purified using the Zymo DNA Clean & Concentrator-5 kit.

C. Negative controls

Field negative controls were collected at each site prior to the collection of field samples. Field blanks were collected by filtering one liter of distilled water through a glass microfiber filter. Field blank filter samples were processed as described above for field samples. Extraction negative controls were also collected, consisting of unused glass microfiber filters, which were processed in parallel with field samples for DNA extraction. During PCR amplification, no template control reactions were included in all sets of PCR reaction. If any no template control reaction showed detectable product on the agarose gel, the entire set of reactions were discarded and rerun.

D. Sequencing

DNA sequencing was performed by Genewiz, Inc with Illumina HiSeq 2x150 bp paired-end (PE) protocols. Barcoded samples were sequenced in a single run with a targeted output of 350 million reads. Prior to sequencing, sample PCR products were pooled, with an equivalent mass of product for each sample added. For field and laboratory blanks that did not show amplification, and equivalent volume of the PCR reaction was added to the pooled mixture.

E. Analysis

1. DNA sequence data processing

DNA sequencing data were initially trimmed and demultiplexed using cutadapt (Martin, 2011) to orient all reads in the forward direction. Further data processing was then performed in R Studio (RStudio Team, 2020) using dada2 (Callahan et al., 2016) to denoise sequences, identify amplified sequence variants (ASVs), and quantify the number of ASVs for each sample.

F. Taxonomic assignment

Taxonomic assignment of ASVs was initially performed using the BLAST using both the full nucleotide collection (nr/nt) collection in the National Center for Biotechnology Information (NCBI; https://blast.ncbi.nlm.nih.gov/) and a custom database of reference species for expected taxa.

Taxonomic assignments were further verified through phylogenetic reconstruction of ASV sequences and reference sequences for all native and non-native fish known to occur in Arizona, along with reference sequences for closely related species.

III. Results and Discussion

A. CAP and Lake Pleasant eDNA sampling – Summer 2022

Sequencing and analysis of samples collected from 12 sites along the CAP canal, Waddell Canal, and Lake Pleasant in August and September 2022 resulted in 249 unique ASVs being identified. Taxonomic assignment of these ASVs identified 164 sequences that matched most closely to fish sequences. Across sites the mean total number of sequence reads matched to fish reference sequences was 9,819,891 reads. The minimum number of sequence reads from a single site was 7,062,740 reads from SALTGILA_US. The maximum number of sequence reads from a single site was 12,782,737 reads from HASS_US.

Based on BLAST hits and phylogenetic reconstruction, these ASVs were found to cluster into 17 distinct groups, interpreted as each corresponding to a single species or genus of origin. The mean number of species detected across sites was 10.8 species. The maximum number of species detected from a single site was 15 species at Lake Pleasant (LAKEPLEA_01). The minimum number of species detected from a single site was 7 species from SALTGILA_US.

1. Grass carp (Ctenopharyngodon idella) and common carp (Cyprinus carpio)

DNA sequence reads matching to grass carp (*C. idella*) and common carp (*C. carpio*) reference sequences were detected in samples from all 12 sites. Sequences for these two species represented the majority of sequences in samples from 6 of the 12 sites sampled (Figure 5 and Figure 6). Grass carp had a maximum percentage of reads per site of 53.6% of reads from SALTGILA_PP. Common carp had a maximum percentage of reads per site of 78.2% from SALTGILA_US. Both species had their minimum percentage of reads from Lake Pleasant (LAKEPLEA_01), with grass carp making up 0.004% of reads and common carp making up 1.9% of reads.

Although sequences matching to grass carp make up only a small percentage of reads from Lake Pleasant, the result is still surprising as there do not appear to be previous reports of grass carp having been stocked or caught in the reservoir. Given that all associated filed and lab controls showed no sign of contamination the detection appears to be valid, although the DNA fragments captured could have been introduced from an exogenous source rather than being shed from a live fish in the reservoir.

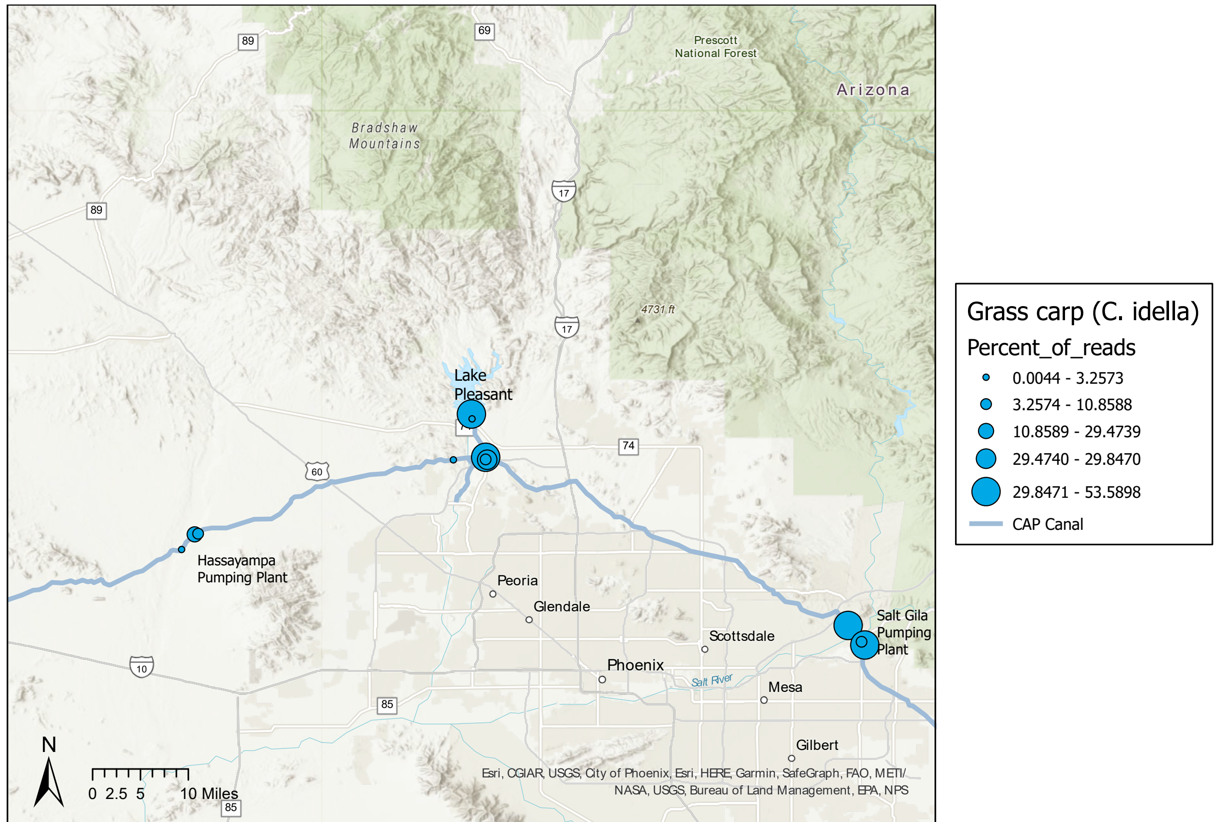

Figure 2: Distribution of grass carp (C. idella) sequence detections.

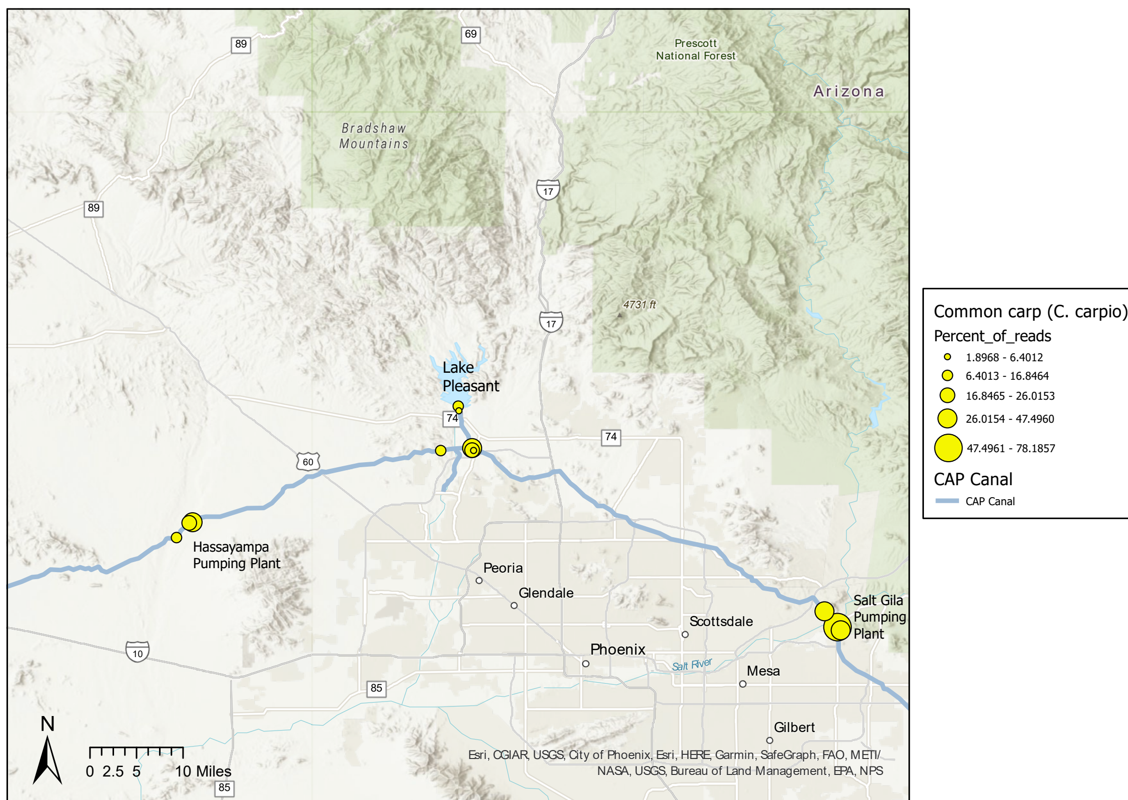

Figure 3: Distribution of common carp (C. carpio) sequence detections.

2. Striped bass (*Morone saxatilis*)

DNA sequences matching to striped bass (*M. saxatilis*) were found in samples from all 12 sites (Figure 4). Read abundance ranged from 0.002% of reads from Lake Pleasant (LAKEPLEA_01) to 13.5% of reads from WAPP_PP. Stiped bass generally showed higher read frequencies in the Waddell Canal and in the CAP Canal below the junction with the Waddell Canal that were observed in samples from the CAP Canal upstream of the junction with the Waddell Canal. Although striped bass represent an important recreational fishery in Lake Pleasant, read frequency was lowest in the samples collected from the reservoir. This may be due to the fact that the sample was collected from the surface during late summer, when temperate striped bass would likely favor cooler waters deeper in the reservoir. This would also explain the high read frequency at WADD_PP, which during the sampling is an outlet site releasing water from deeper in Lake Pleasant.

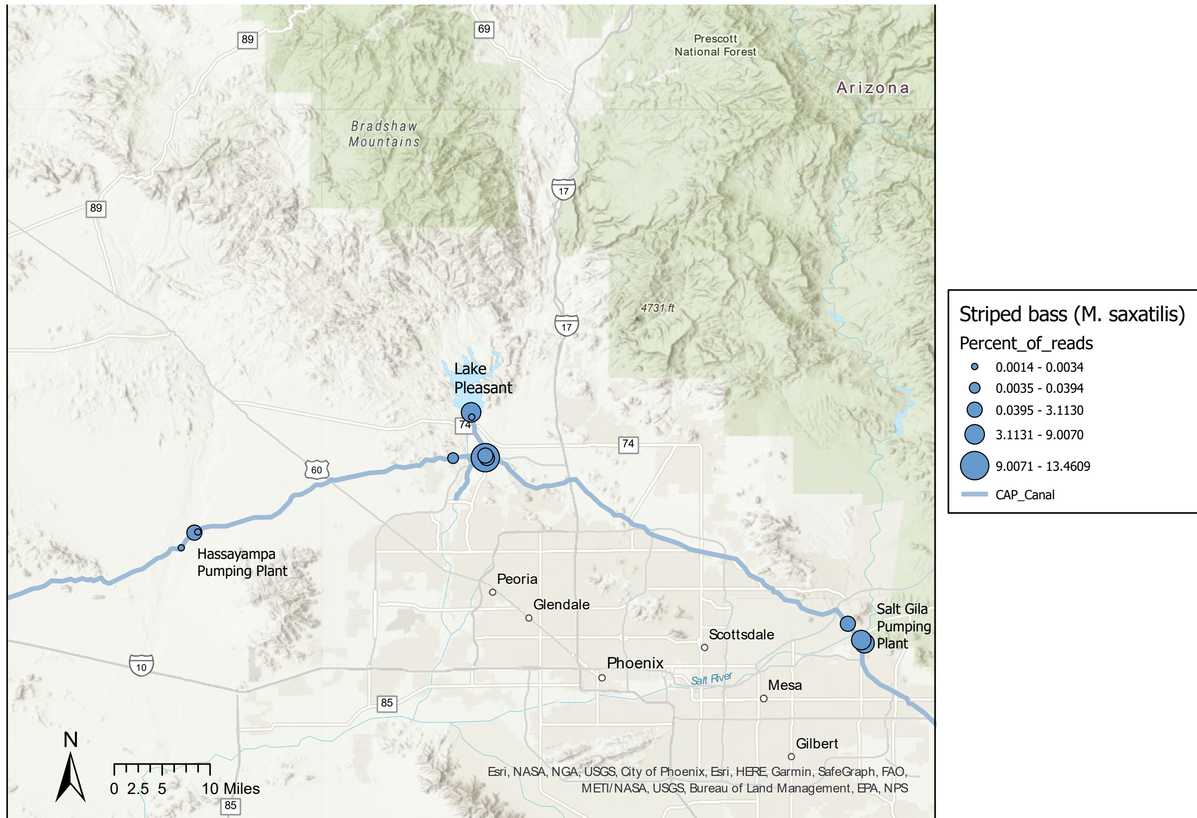

Figure 4: Distribution of striped bass (M. saxatilis) sequence detections.

3. Channel catfish (*Ictalurus punctatus*) and Blue catfish (*Ictalurus furcatus*)

DNA sequences matching to channel catfish (*I. punctatus*) were detected in samples from all 12 sites (Figure 5). Read frequency ranged from 0.0005% of reads from Lake Pleasant (LAKEPLEA_01) to 2.76% of reads from SALTGILA_DS. As with striped bass, the read frequency was quite lowest in Lake Pleasant but near the top of the range at WADD_PP immediately downstream.

DNA sequences matching to the congeneric blue catfish (*I. furcatus*) were detected from only one site, WASG_02, where they represented 0.008% of reads (Figure 6).

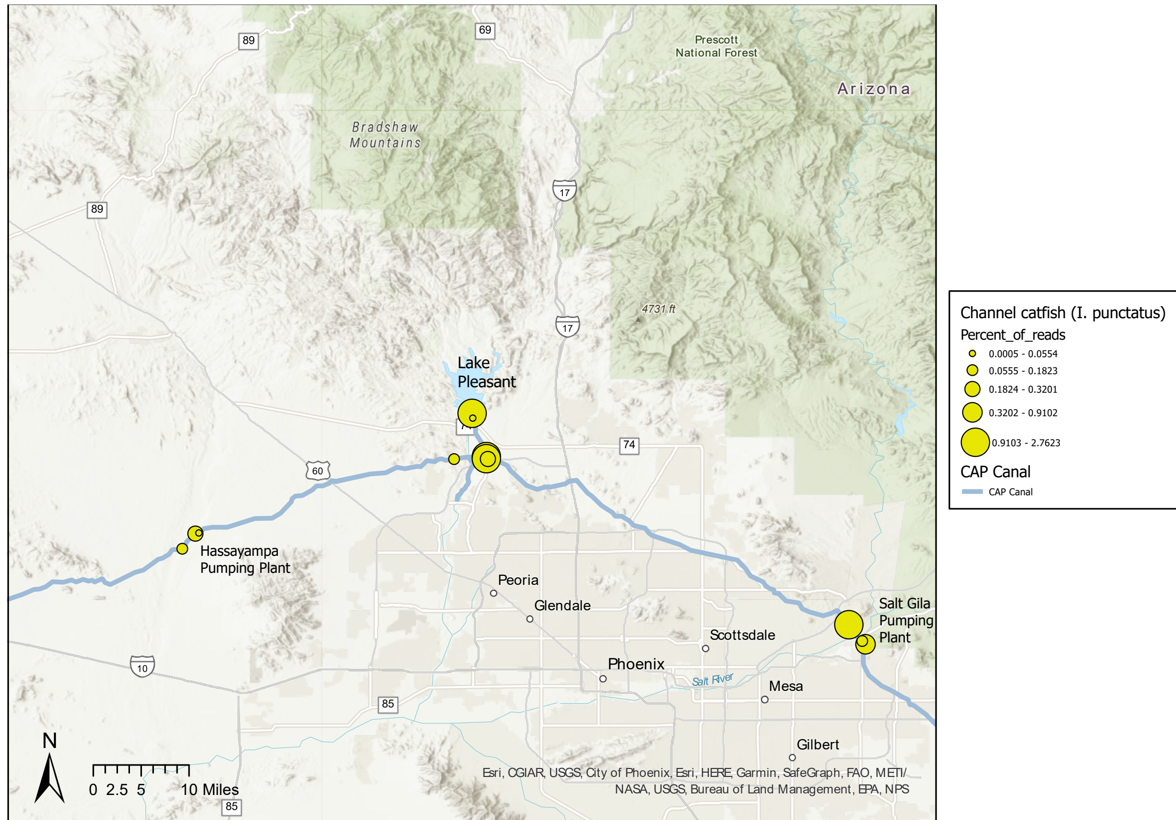

Figure 5: Distribution of channel catfish (I. punctatus) sequence detections.

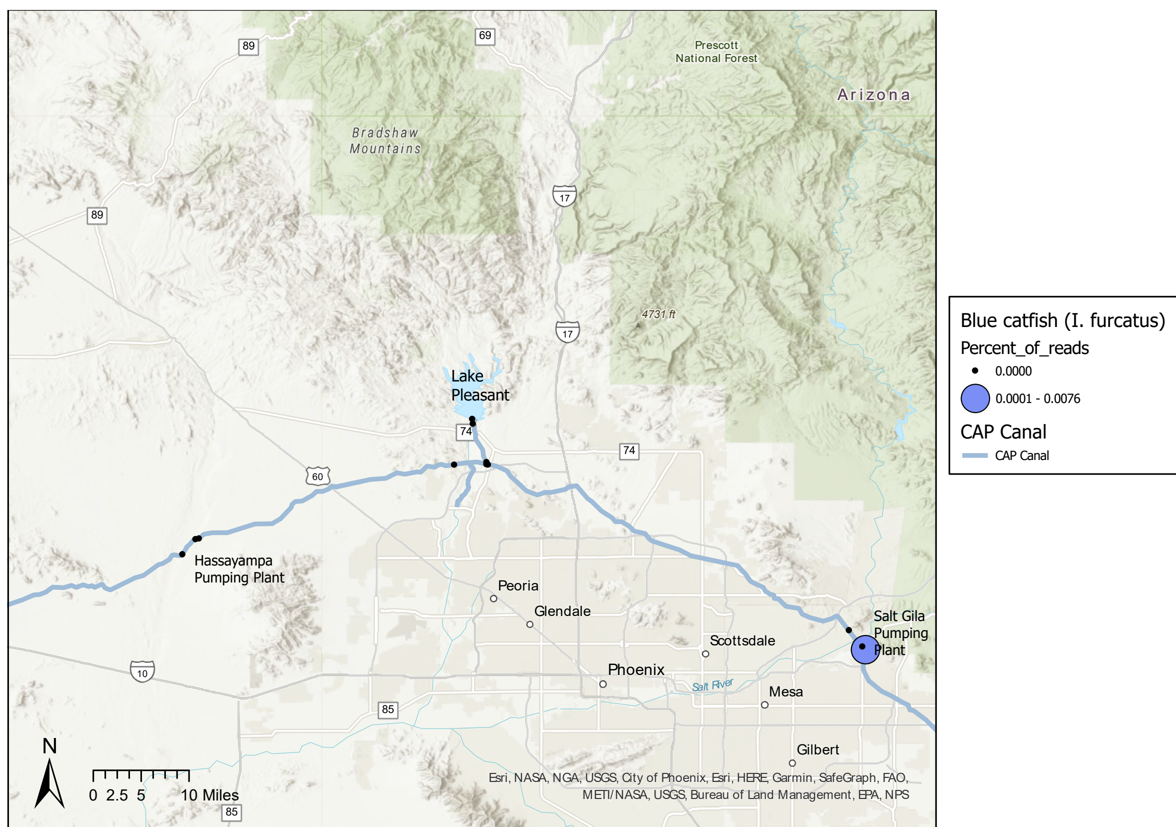

Figure 6: Distribution of blue catfish (I. furcatus) sequence detections.

4. Green sunfish (*Lepomis cyanellus*), Bluegill (*Lepomis macrochirus*), and Redear sunfish (*Lepomis microlophus*)

DNA sequences for three distinct species of sunfish in the genus *Lepomis*, green sunfish (*L. cyanellus*), bluegill (*L. macrochirus*), and redear sunfish (*L. microlophus*), were detected from the samples.

DNA sequences matching to green sunfish were detected in samples from all 12 sites (Figure 7). Sequence read frequencies for green sunfish ranged from 0.002% from SALTGILA_PP to 53.2% from WADD_01.

DNA sequences matching to bluegill were detected in samples from 11 sites (Figure 8). Read frequencies ranged from 0.0002% from WASG_02 to 25.6% from WADD_PP. No reads matching green sunfish were detected in samples from SALTGILA_PP.

DNA sequences matching to redear sunfish were detected in samples from 3 sites, including Lake Pleasant and two sites upstream of the Waddell Canal (Figure 9). Read frequencies for

redear sunfish ranged from 0.0004% from HASS_US to 5.02% from Lake Pleasant (LAKEPLEA_01).

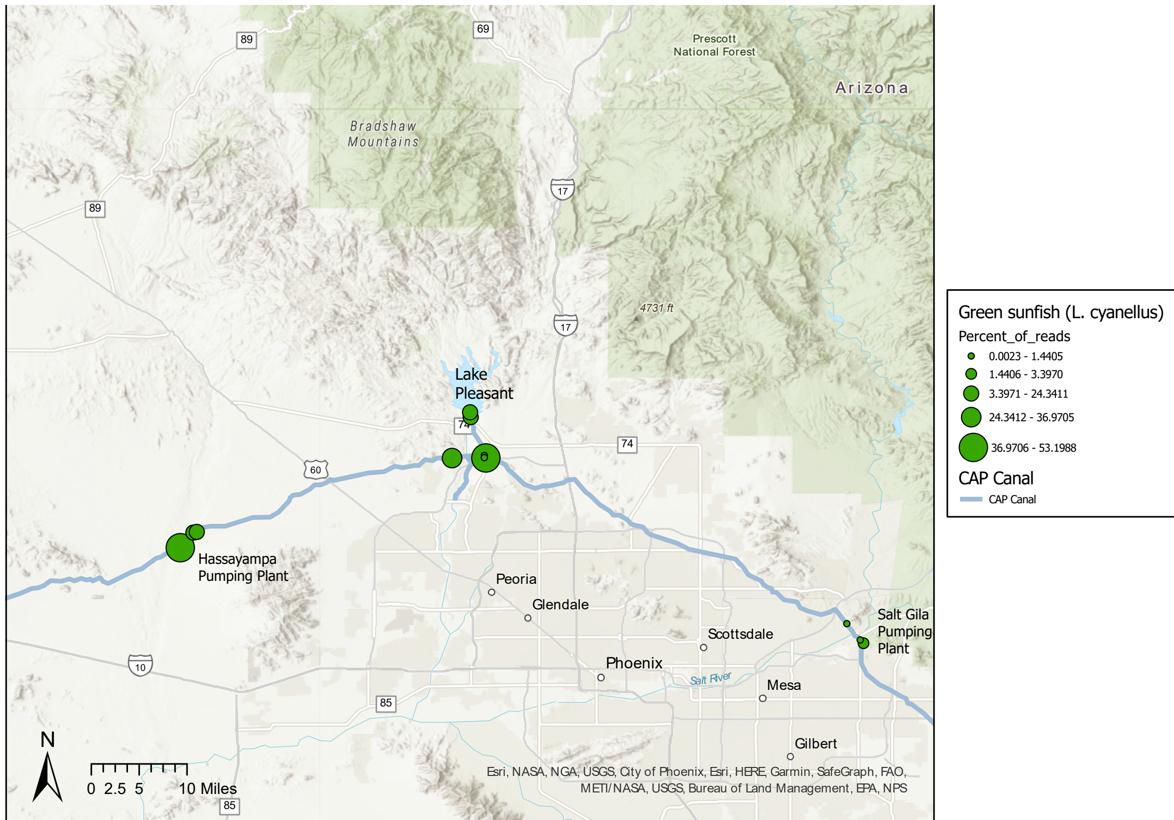

Figure 7: Distribution of green sunfish (L. cyanellus) sequence detections.

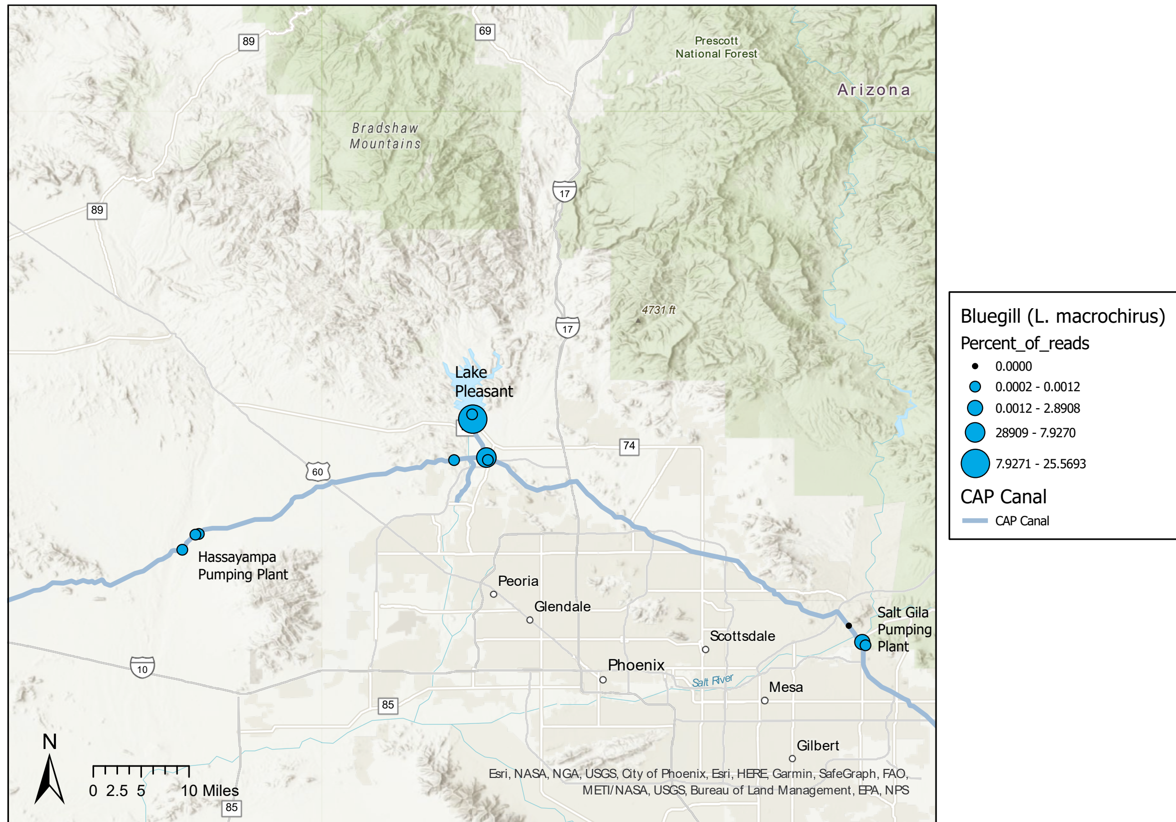

Figure 8: Distribution of bluegill (L. macrochirus) sequence detections.

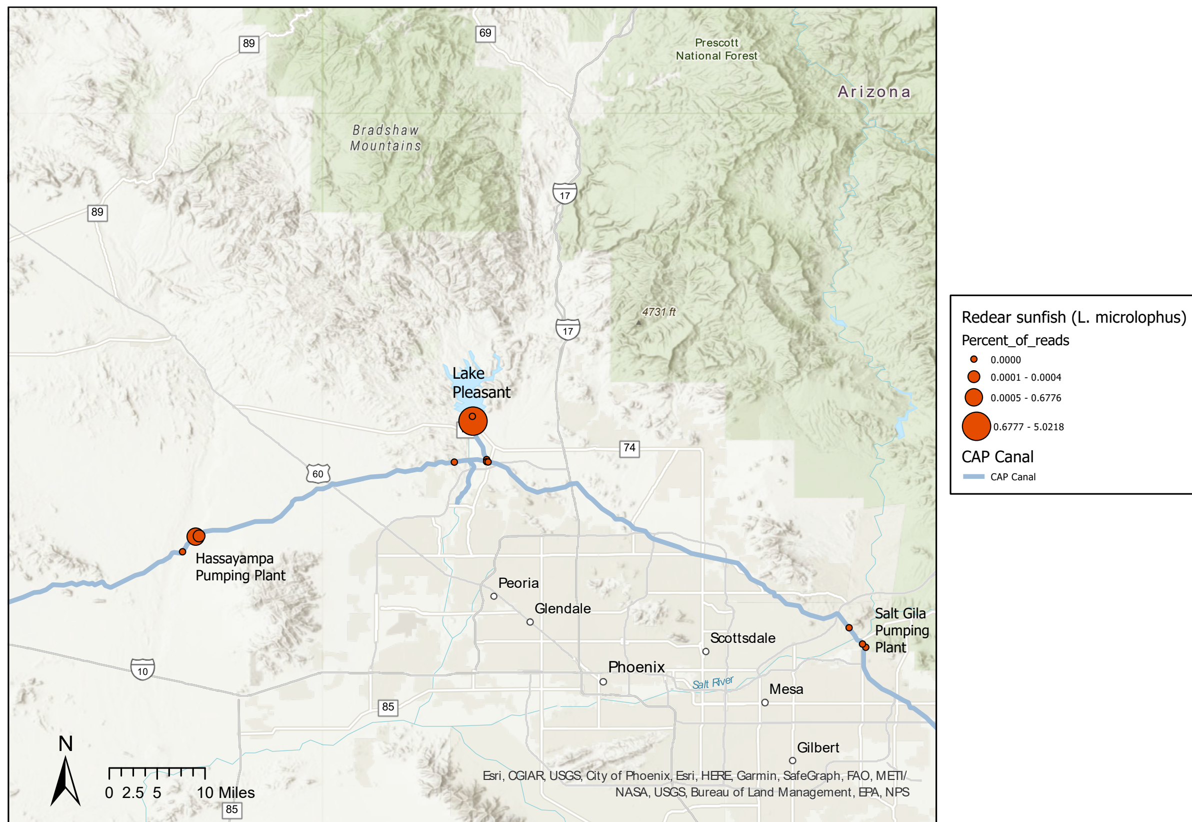

Figure 9: Distribution of redear sunfish (L. microlophus) sequence detections.

5. Inland silverside (*Menidia beryllina*)

Sequences matching to inland silverside (*M. beryllina*) were detected at 11 of the sites sampled, with SALTGILA_US being the one site with no matching sequences detected (Figure 10). Read frequencies ranged from 0.0002% from SALTGILA_PP to 35.7% from Lake Pleasant (LAKEPLEA_01). Notably, the prevalence of inland silverside detections contrasts with our previous study where no detections were made from samples collected throughout the CAP Canal (Passamaneck 2022).

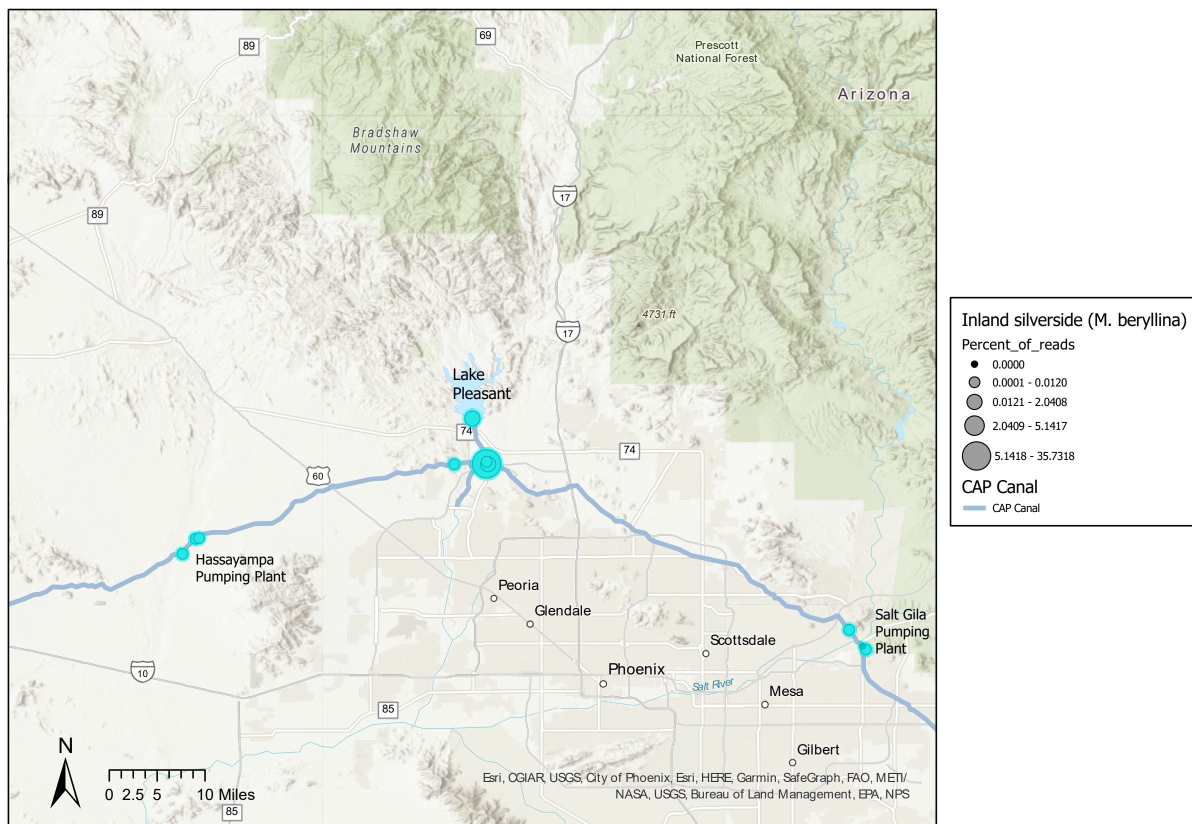

Figure 10: Distribution of inland silverside (M. beryllina) sequence detections.

6. Smallmouth bass (*Micropterus dolomieu*) and Largemouth bass (*Micropterus salmoides*)

Sequences matching to two species of the genus *Micropterus*, smallmouth bass (*M. dolomieu*) and largemouth bass (*M. salmoides*) were detected in samples.

Sequences matching to smallmouth bass were detected in samples from 6 sites (Figure 11). Read frequencies ranged from 0.0005% from WADD_01 to 8.1% from WASG_01.

Sequences matching to largemouth bass were detected only in samples from Lake Pleasant (LAKEPLEA_01) where they represented 1.86% of reads (Figure 12).

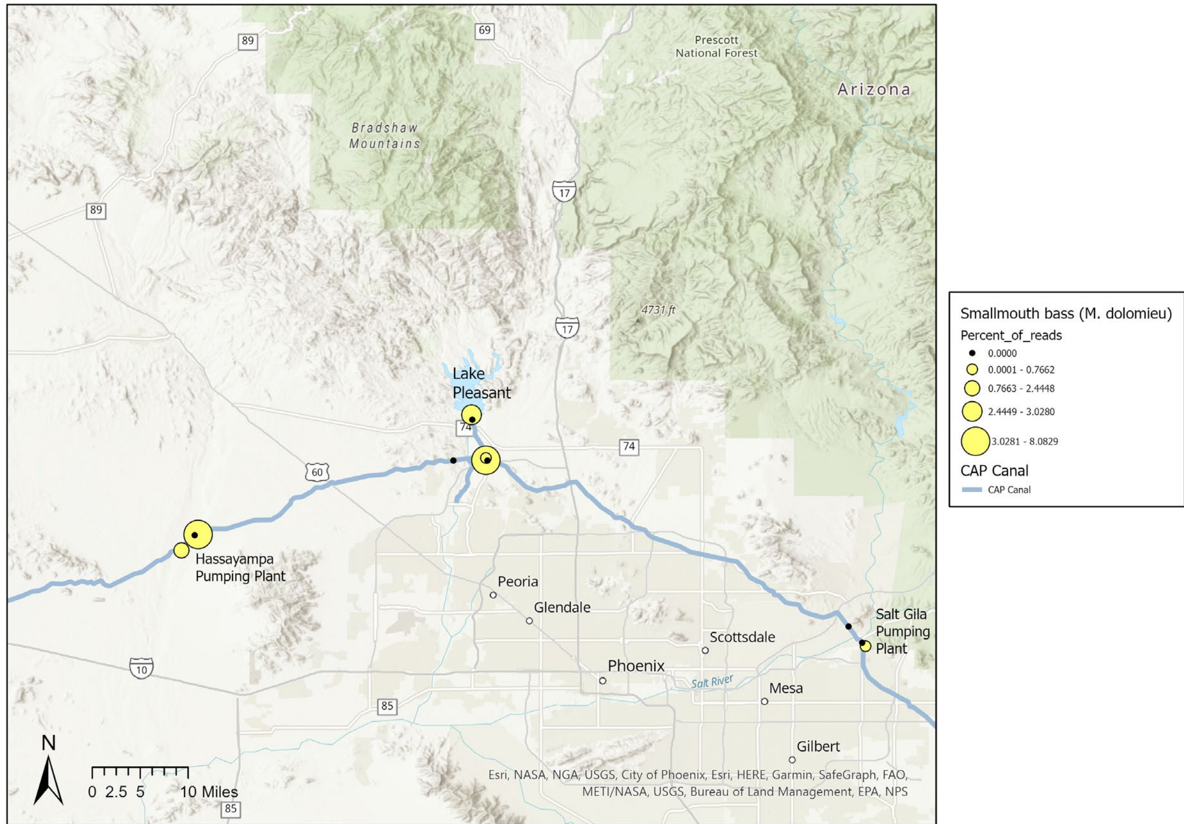

Figure 11: Distribution of smallmouth bass (M. dolomieu) sequence detections.

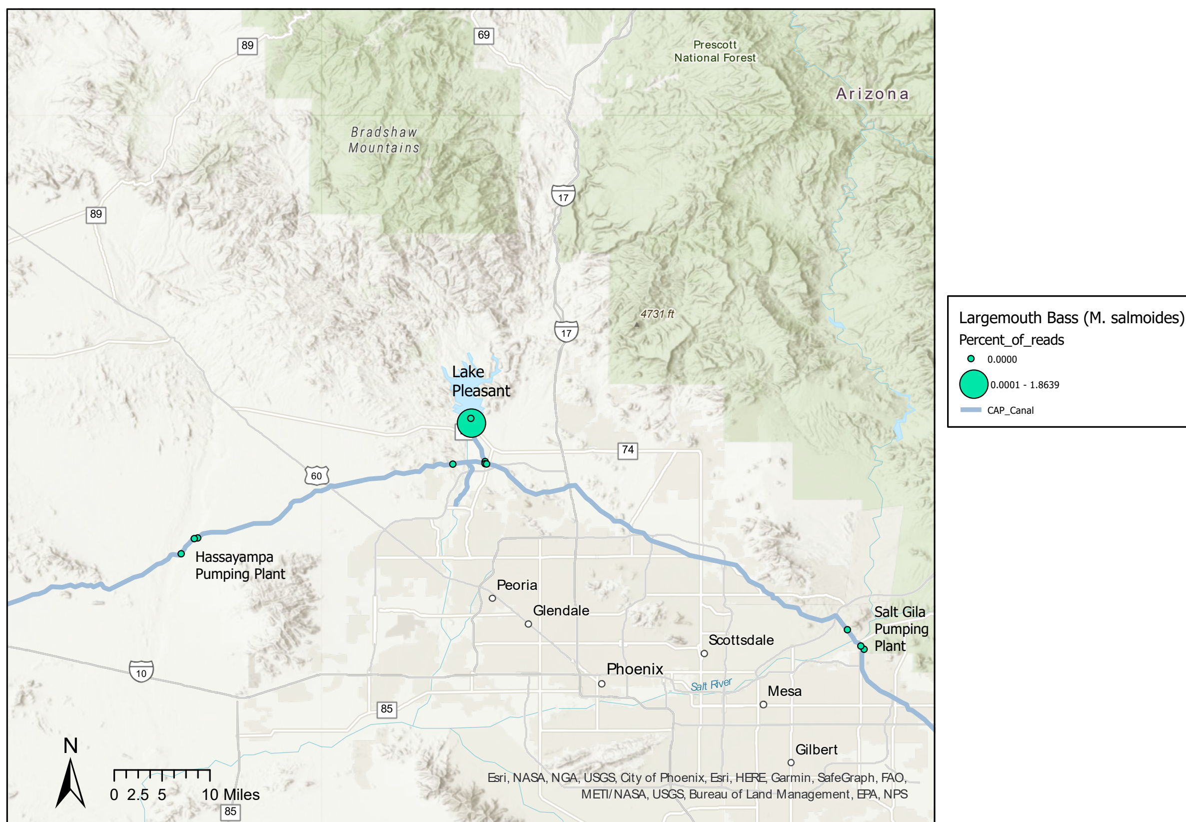

Figure 12: Distribution of largemouth bass (M. salmoides) sequence detections.

7. Gizzard shad (*Dorosoma cepedianum*) and Threadfin shad (*Dorosoma petenense*)

Sequences matching to two species of the genus *Dorosoma*, gizzard shad (*D. cepedianum*) and threadfin shad (*D. petenense*), were detected in samples.

Sequences matching to gizzard shad (*D. cepedianum*) were detected in samples from 9 sites (Figure 13). Read frequencies ranged from 0.0005% from HASS_PP and WASG_01 to 21.99% from Lake Pleasant (LAKEPLEA_01).

Sequences matching to threadfin shad (*D. petenense*) were detected in samples from 10 sites (Figure 14). Read frequencies ranged from 0.0003% from WASG_01 and SALTGILA_DS to 14.0% from HASS_US.

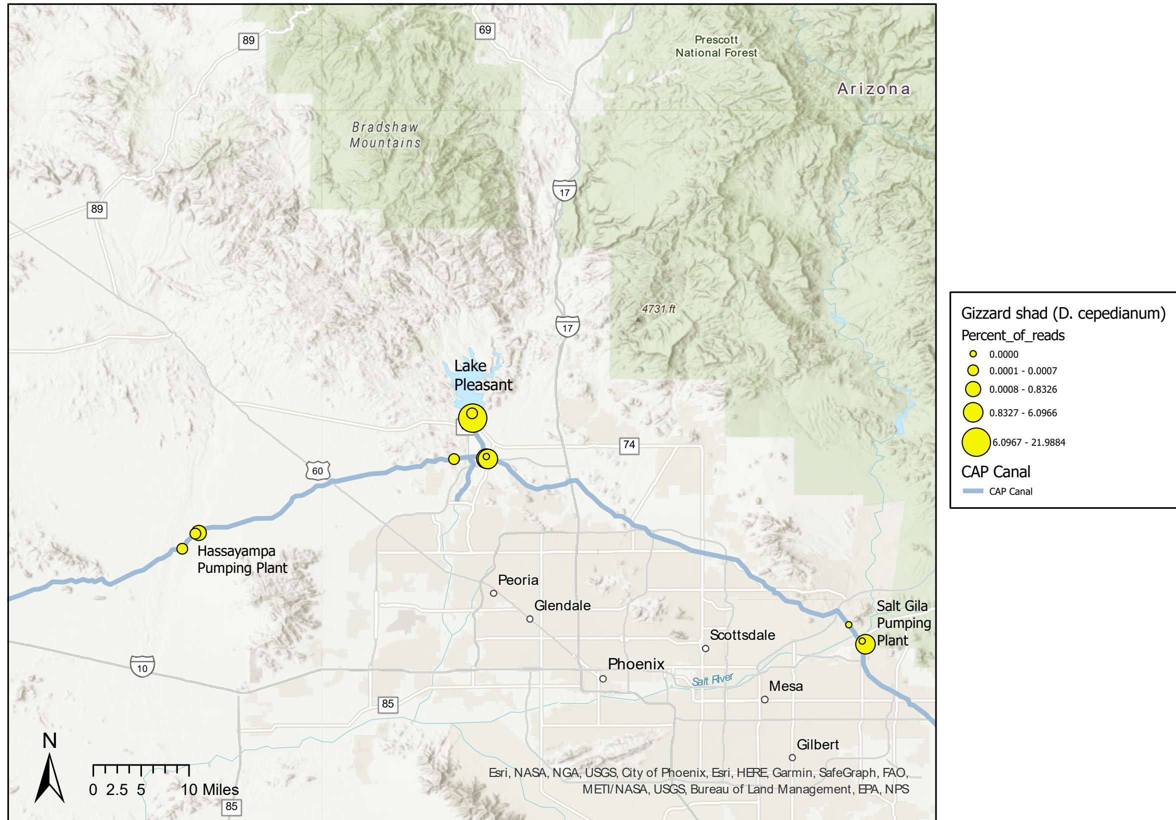

Figure 13: Distribution of gizzard shad (D. cepedianum) sequence detections.

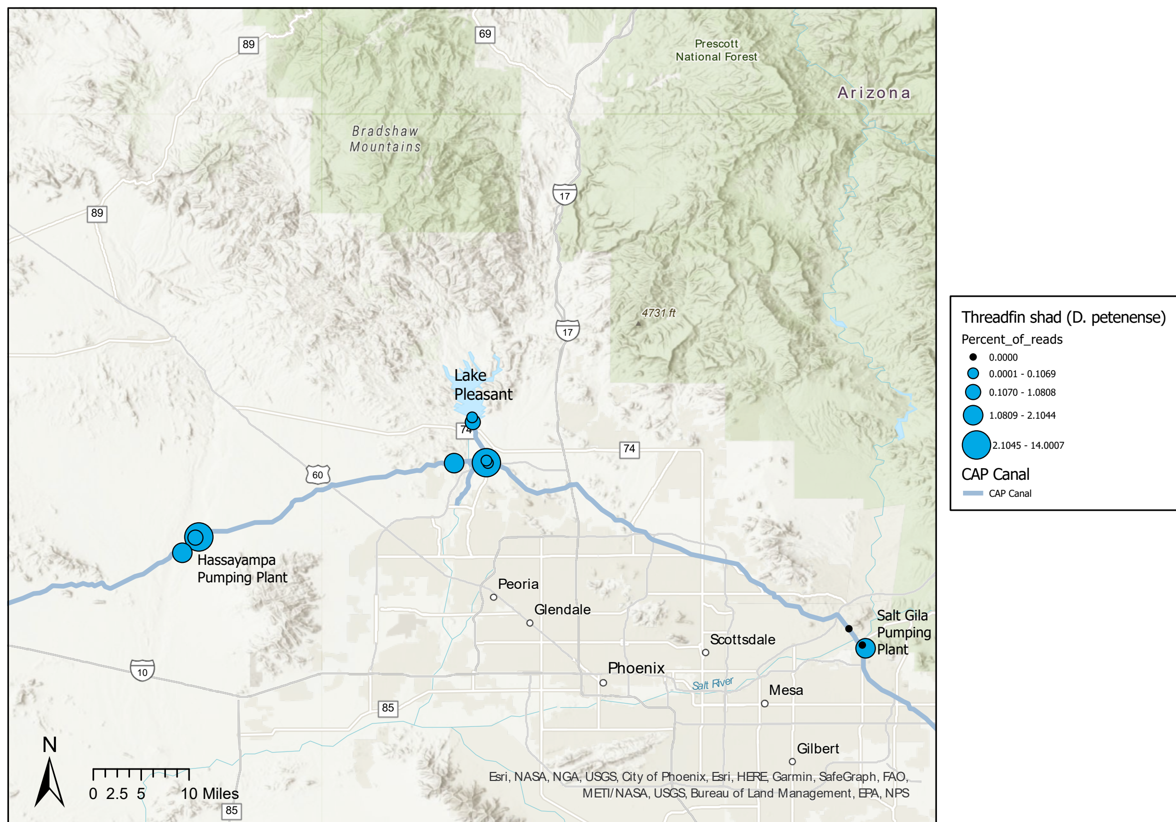

Figure 14: Distribution of threadfin shad (D. petenense) sequence detections.

8. Western mosquitofish (*Gambusia affinis*)

Sequences matching to Western mosquitofish (*G. affinis*) were detected in samples from all 12 sites (Figure 15). Read abundances ranged from 0.0011% from SALTGILA_PP to 41.99% from HAWA_06. Read frequencies for mosquitofish were notably higher in samples from upstream of the Waddell Canal than they were from downstream of it, or in the Waddell Canal itself and Lake Pleasant.

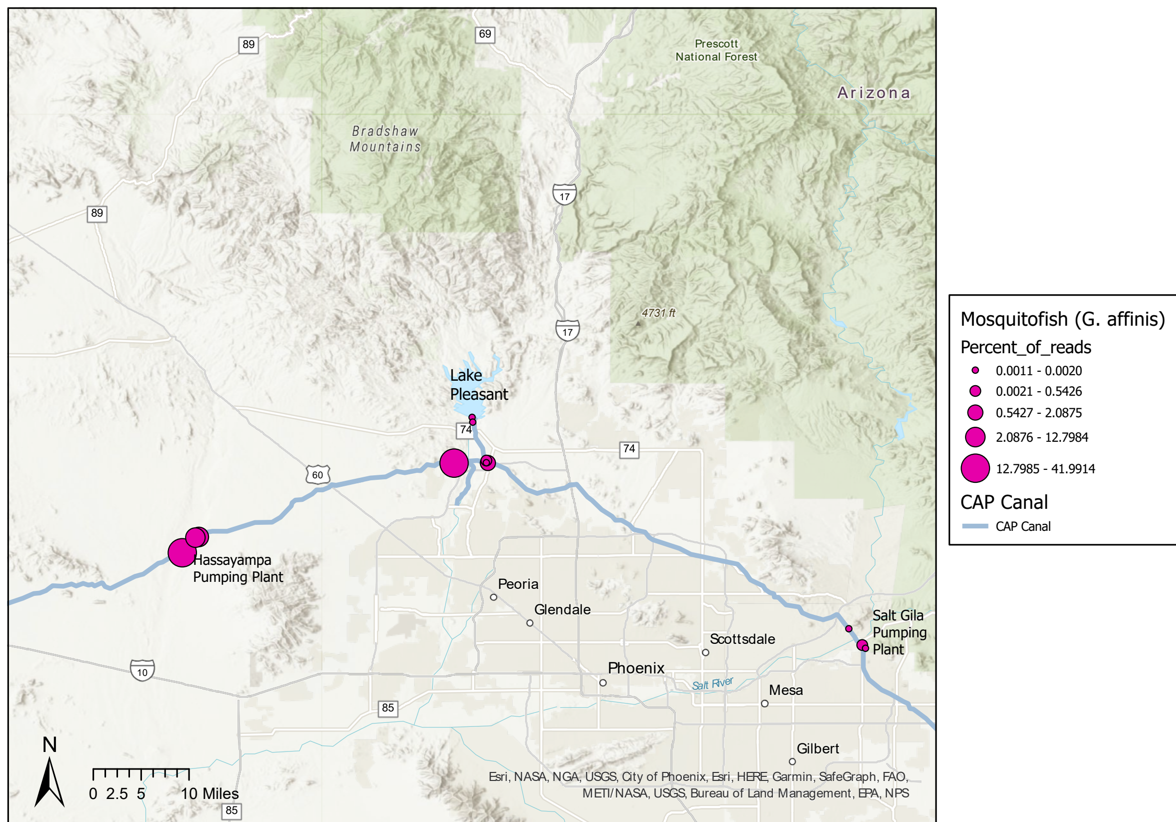

Figure 15: Distribution of Western mosquitofish (G. affinis) sequence detections.

9. Tilapia (*Oreochromis* sp.)

Sequences matching to tilapia (*Oreochromis* sp.) were detected from Lake Pleasant (LAKEPLEA_01) but were not detected in samples from any other sites (Figure 16). Reads matching to tilapia made up 0.66% of the reads from Lake Pleasant. Sequences for tilapia were represented by a single ASV, which had a 100% identity match to reference sequences for both blue tilapia (*O. aureus*) and Nile tilapia (*O. niloticus*). For this reason, the detection is matched to the genus level (*Oreochromis* sp.) rather than to a specific species.

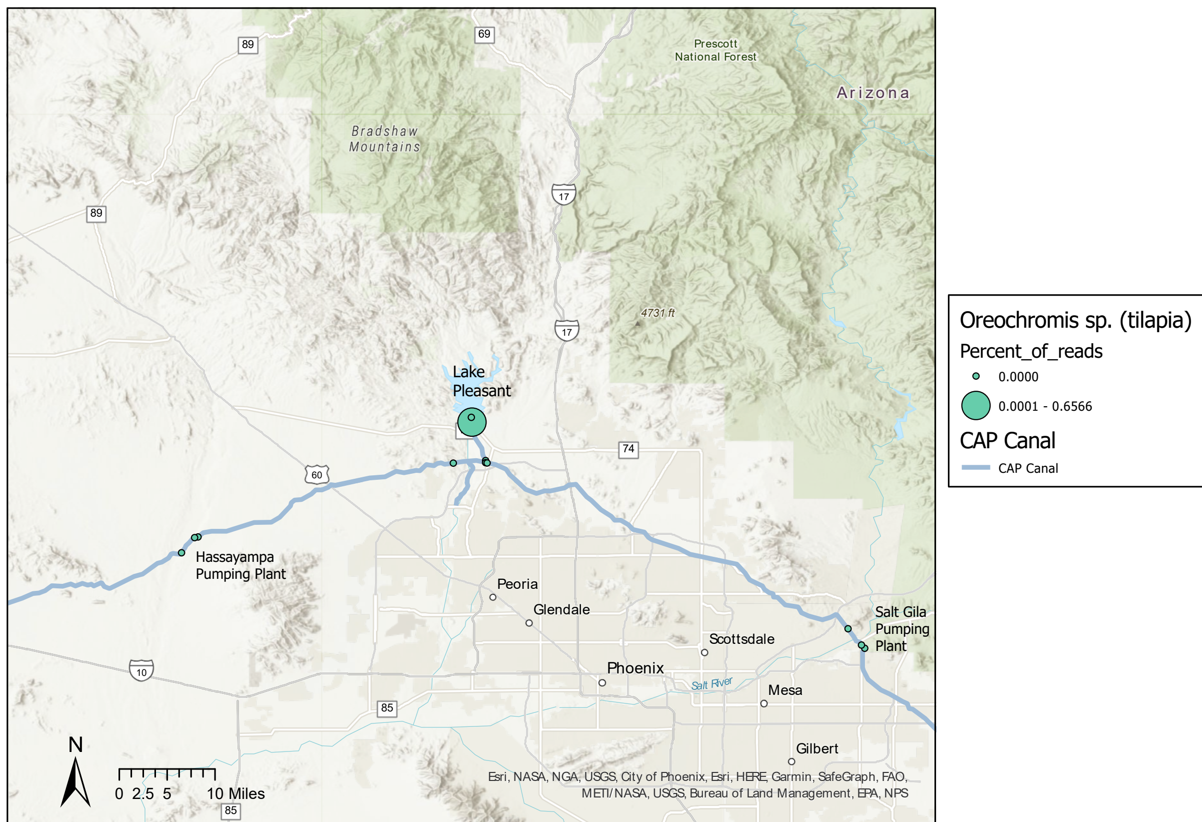

Figure 16: Distribution of tilapia (Oreochromis sp.) sequence detection.

10. Flathead catfish (*Pylodictis olivaris*)

Along with the channel catfish (*I. punctatus*) and blue catfish (*I. furcatus*) discussed above, DNA sequences matching to flathead catfish (*P. olivaris*) were also detected at four sites, including three sites upstream of Waddell Canal and from Lake Pleasant (LAKEPLEA_01) (Figure 17). Read frequencies for sequences matching to flathead catfish ranged from 0.0015% from HAWA_06 to 0.04% from Lake Pleasant. Matching sequences were not detected in samples from Waddell Canal or the CAP Canal downstream of the junction with Waddell Canal.

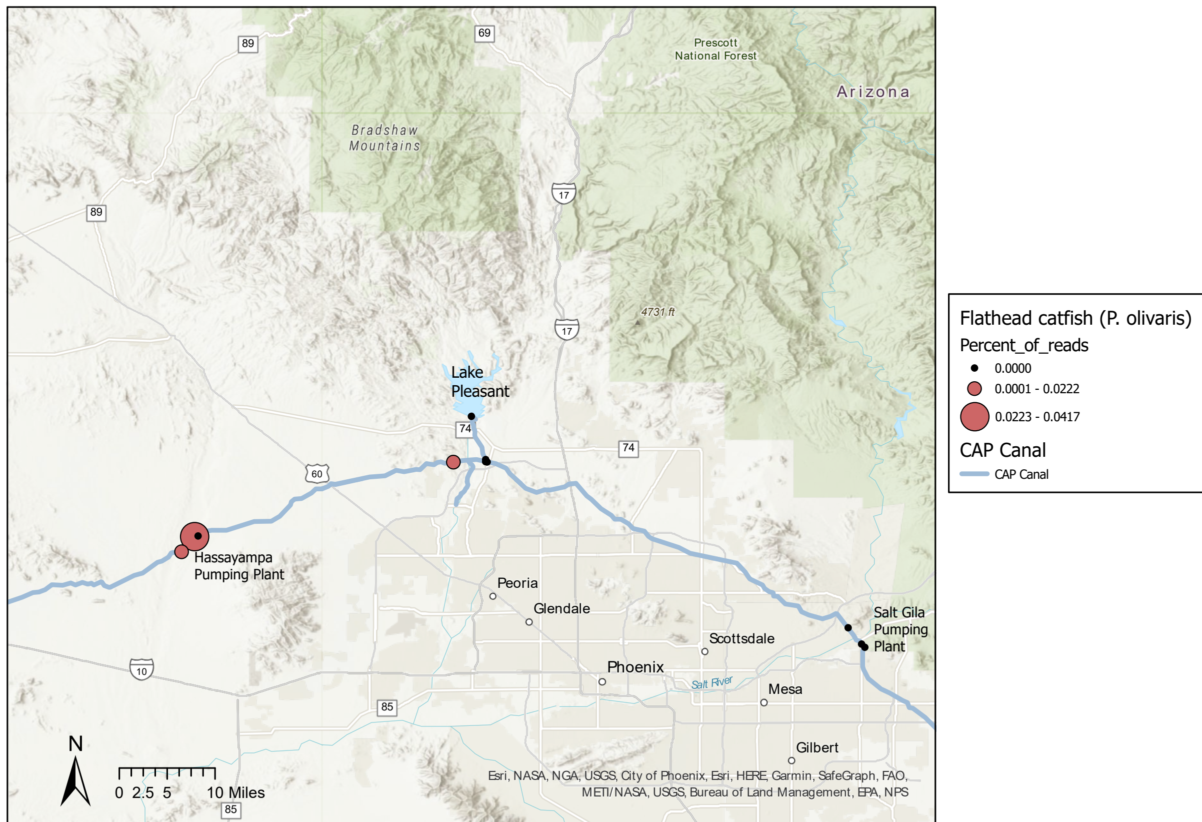

Figure 17: Distribution of flathead catfish (P. olivaris) sequence detections.

11. Salmon/trout (*Oncorhynchus* spp.)

Sequences matching to salmon and trout (*Oncorhynchus* spp.) were detected from a single site, Salt Gila Pumping Plant (SALTGILA_PP) (Figure 18). These sequences comprised 0.77% of the reads from Salt Gila Pumping Plant. The single ASV identified was most similar to reference sequences for chum salmon (*O. keta*) with 99.41% identity, but also had greater than 98% identity to reference sequences from sockeye salmon (*O. nerka*), coho salmon (*O. kisutch*), and rainbow trout (*O. mykiss*). For this reason, identification is made only to the level of genus, rather than to a specific species. The source of these sequences is uncertain. No trout, kokanee, or other salmonids have been reported from the CAP Canal or Lake Pleasant, although rainbow trout are reported to be established in the Salt River, Verde River, and Arizona Canal (nas.er.usgs.gov). In our previous study of fish eDNA from the CAP Canal we also detected sequences matching to *Oncorhynchus* spp. from a site 51.5 canal miles downstream of Salt Gila Pumping Plant (SGBR_012) (Passamaneck, 2022). Comparison of these two ASVs shows that they have 98.2% identity, with three positions having substitutions or indels. This suggests that the two sequences are likely from different sources. Given that no sequence was detected in associated filed or laboratory controls it appears that the detection is valid. However, whether the source of the sequences was a fish in the canal system or an exogenous source cannot be determined.

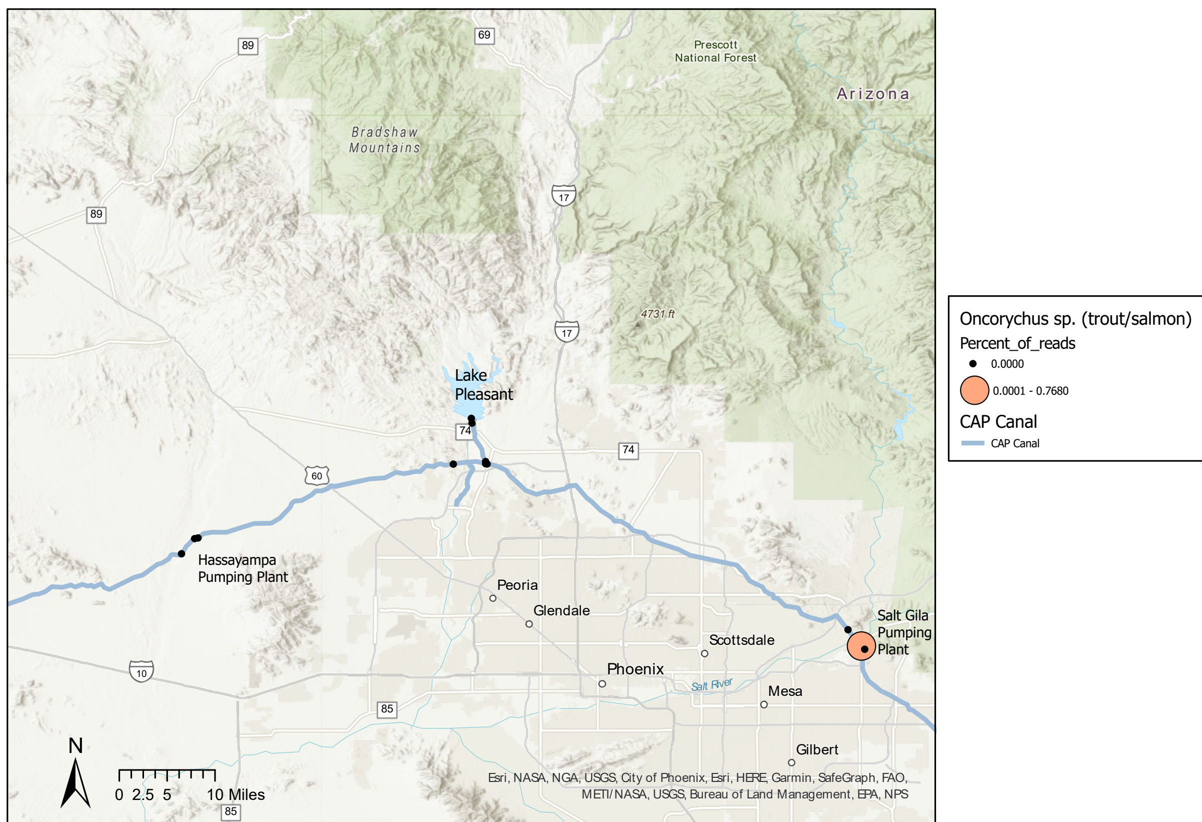

Figure 18: Distribution of salmon/trout (Oncorhynchus spp.) sequence detection.

12. Non-fish sequences

Although the MiFish primers are designed to be specific to fish, they did display some cross-reactivity with non-fish species. Eighty-five of the 249 recovered ASVs did not to match to any available reference sequence for fish species. These non-fish ASVs accounted for 3.6% of the total reads. These sequences were analyzed separately by BLAST search against the GenBank nr/nt databases. The majority of these non-fish ASVs had fewer than 20 reads per sample (the threshold for retention) and were discarded without further analysis. Twelve of the non-fish ASVs had 20 or more reads in at least one sample. Seven of these ASVs matched to reference sequences for mammals, including bats, pig, cow, and human, and one matched to mallard duck. Four of the non-fish ASVs had matches below 98% to reference sequences in the GenBank nr/nt databases and may have been the results of chimeras or other errors in sequencing library preparation.

B. Distributions in Lake Pleasant, Waddell Canal and the CAP Canal

In total sequences for 15 of the 17 species/genera of fish detected in this study were present in samples from Lake Pleasant. The detected taxa fish, with the exception of grass carp (discussed above), have all previously been caught in Lake Pleasant by traditional surveys (Clarkson et al., 2011, Gill et al, 2019, Stewert et al., 2007). Two of the fish taxa detected, largemouth bass (*M. salmoides*) and tilapia (*Oreochromis* spp.) were found exclusively in samples from Lake Pleasant. Both these fish taxa were detected in samples from the CAP Canal during our previous study (Passamaneck, 2022). Several species previously identified from Lake Pleasant were not detected from our eDNA samples, including white bass (*M. chrysops*), white crappie (*P. annularis*), black crappie (*P. nigromaculatus*), red shiner (*Cyprinella lutrensis*), golden shiner (*Notemigonus crysoeucas*), and yellow bullhead (*Ameiurus natalis*). The absence of detection of black crappie in samples from Lake Pleasant or any other site is notable given that it was detected at every site sampled along the CAP Canal in our previous study, albeit at low frequencies (Passamaneck, 2022). The absence of white crappie detection is consistent with reports that this species has been caught at low levels by traditional surveys over the last two decades (L Pleasant Management Plan). A recent study which conducted extensive gill net and electrofishing surveys of Lake Pleasant also failed to capture any white crappie (Gill and Jones 2019; Joshua Grant, personal communication).

Species detections and read frequencies differed considerably between Lake Pleasant (LAKEPLEA_01) and the two sites Waddell Canal (WADD_PP and WADD_01) (Figure 19). Reads for gizzard shad, green sunfish, and inland silverside were most numerous in the samples from Lake Pleasant, and were detected at much lower frequencies from WADD_PP. Reads for common carp, threadfin shad, bluegill, and striped bass were most numerous from WAD_PP, while reads for grass carp and green sunfish were most numerous from WADD_01. As discussed above, variance in read frequencies between LAKEPLEA_01 samples and WADD_PP might reflect difference in habit usage, with LAKEPLEA_01 samples capturing DNA from species near the surface of Lake Pleasant and WADD_PP samples reflecting populations deeper in the lake near the intakes for the for the Waddell Pump/Generating Plant. Alternatively, the differences between LAKEPLEA_01 and WADD_PP may limit export of eDNA from Lake Pleasant, with reads from WADD_PP coming from population in immediate proximity to the sampling site.

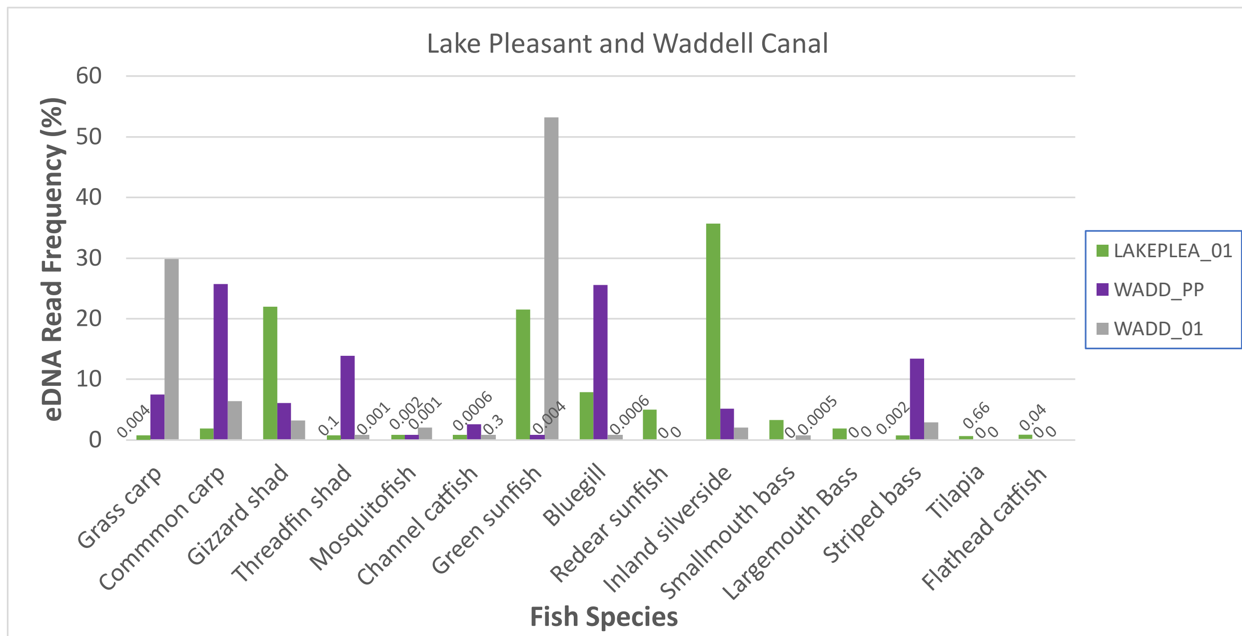

Figure 19: Percent read frequencies for fish species detected from eDNA metabarcoding from LAKEPLEA_01, WADD_PP, and WADD_01. Reads from LAKEPLEA_01 are displayed in green. Reads from WADD_PP are shown in purple. Reads from WADD_01 are shown in grey. For species where the read frequency was below 1%, the value is displayed above bar.

C. Seasonal differences in eDNA

Eight of the sites along the CAP canal that were surveyed for fish eDNA during the summer of 2022 were previously surveyed in February 2021 using the same methods (Passamaneck, 2022). This resampling allows for comparisons between seasons and during different flow regimes. Sampling for the current study was conducted during the summer when water was being released from Lake Pleasant into Waddell Canal, which then fed into the CAP Canal at site WASG_01. During the previous study sampling was conducted in the winter when water is pumped from the CAP canal into Lake Pleasant, and there should have been no contribution of eDNA from fish in Lake Pleasant to samples collected in the CAP Canal.

1. Upstream of Waddell Canal and Lake Pleasant

Four sites upstream of the confluence of the CAP Canal and Waddell Canal were sampled in both Winter 2021 and Summer 2022 for fish eDNA metabarcoding. This included HASS_US (Figure 20), HASS_PP (Figure 21), HASS_DS (Figure 22), and HAWA_06 (Figure 23). From all the upstream sites reads matching to grass carp and common carp made up a smaller proportion of reads in Summer 2022 than in Winter 2021. Conversely, reads matching to green sunfish were a detected at higher frequency in Summer 2022 than in Winter 2021. Mosquitofish, which was not detected in samples collected in Winter 2021 from these sites, had a read frequency between 8.3% and 42% in samples from Summer 2022. All upstream samples from Summer 2022 also had more total species detected than those from Winter 2021, with anaverage of 11.25 species for site in Summer 2022 versus 5.75 species per site in Winter 2021.

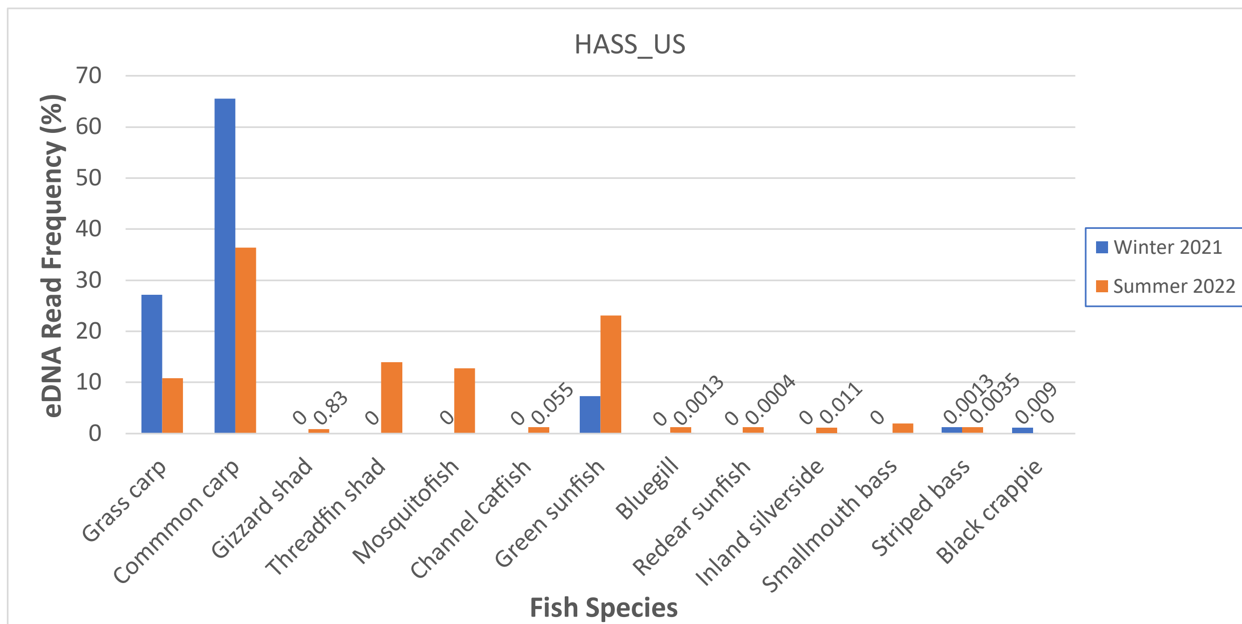

Figure 20: Percent read frequencies for fish species detected from eDNA metabarcoding from HASS_US. Reads from samples collected in the winter of 2021 are displayed in blue. Reads from samples collected in the summer of 2022 are shown in orange. For species where the read frequency was below 1%, the value is displayed above bar.

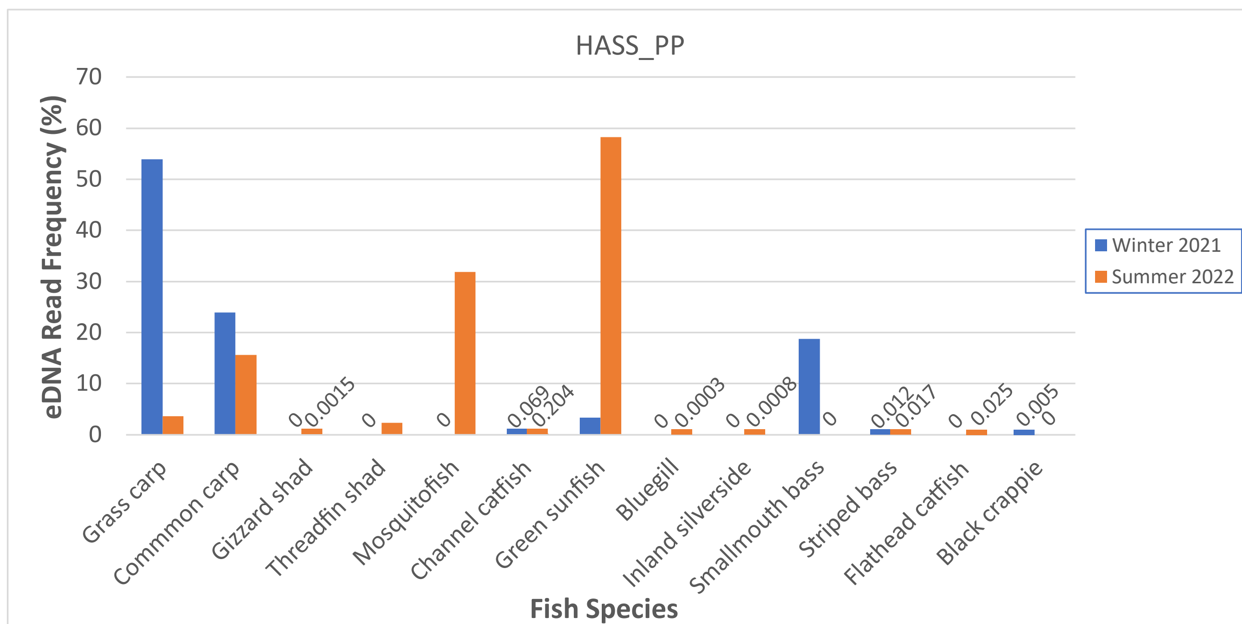

Figure 21: Percent read frequencies for fish species detected from eDNA metabarcoding from HASS_PP. Reads from samples collected in the winter of 2021 are displayed in blue. Reads from samples collected in the summer of 2022 are shown in orange. For species where the read frequency was below 1%, the value is displayed above bar.

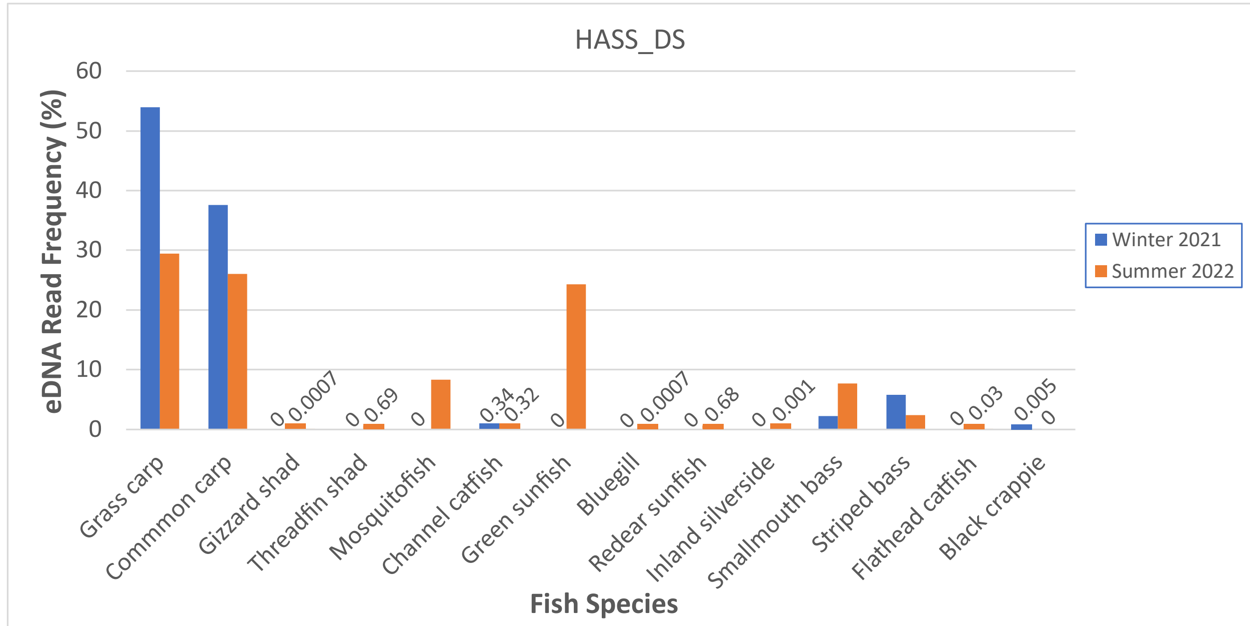

Figure 22: Percent read frequencies for fish species detected from eDNA metabarcoding from HASS_DS. Reads from samples collected in the winter of 2021 are displayed in blue. Reads from samples collected in the summer of 2022 are shown in orange. For species where the read frequency was below 1%, the value is displayed above bar.

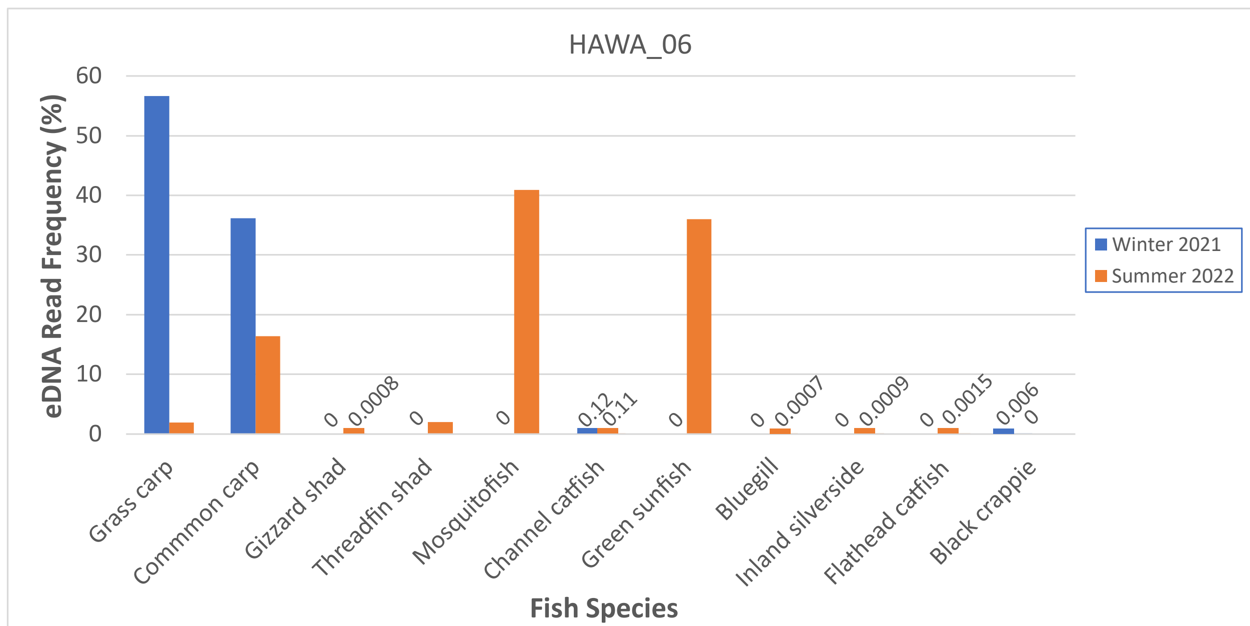

Figure 23: Percent read frequencies for fish species detected from eDNA metabarcoding from HAWA_06. Reads from samples collected in the winter of 2021 are displayed in blue. Reads from samples collected in the summer of 2022 are shown in orange. For species where the read frequency was below 1%, the value is displayed above bar.

2. Downstream of Waddell Canal and Lake Pleasant

Four sites at or downstream of the confluence of the CAP Canal and Waddell Canal were sampled in both Winter 2021 and Summer 2022 for fish eDNA metabarcoding. This included WASG_01 (Figure 24), SALTGILA_US (Figure 25), SALTGILA_PP (Figure 26), and SALTGILA_DS (Figure 27). The pattern of detections from WASG_01 was comparable to those from upstream sites, with the frequency of both grass carp and common carp reads decreasing between Winter 2021 and Summer 2022, while green sunfish, which was not detected in Winter 2021, constituted 20.7% of the reads from Summer 2022. Samples from SALTGILA_US showed relatively little change in the read frequencies for grass carp and common carp between Winter 2021 and Summer 2022, while there was a decrease in the proportion of reads for threadfin shad and mosquitofish, and an increase in the proportion of reads for striped bass. Samples from SALTGILA_PP and SALTGILA_DS both showed a marked increase from Winter 2021 to Summer 2022 in the proportion of reads for grass carp, while the proportion of reads for common carp decreased. At SALTGILA_PP the proportion of striped bass reads decreased between Winter 2021 and Summer 2022. At SALTGILA_DS the proportion of mosquitofish, bluegill, and striped bass reads all decreased between Winter 2021 and Summer 2022.

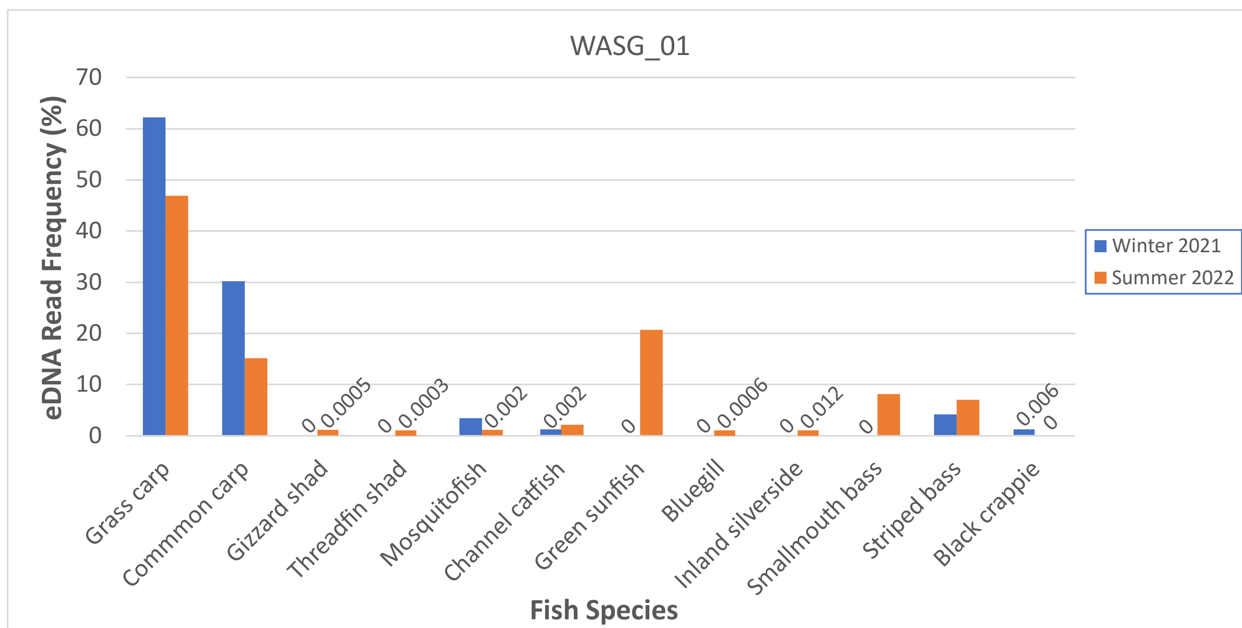

Figure 24: Percent read frequencies for fish species detected from eDNA metabarcoding from WASG_01. Reads from samples collected in the winter of 2021 are displayed in blue. Reads from samples collected in the summer of 2022 are shown in orange. For species where the read frequency was below 1%, the value is displayed above bar.

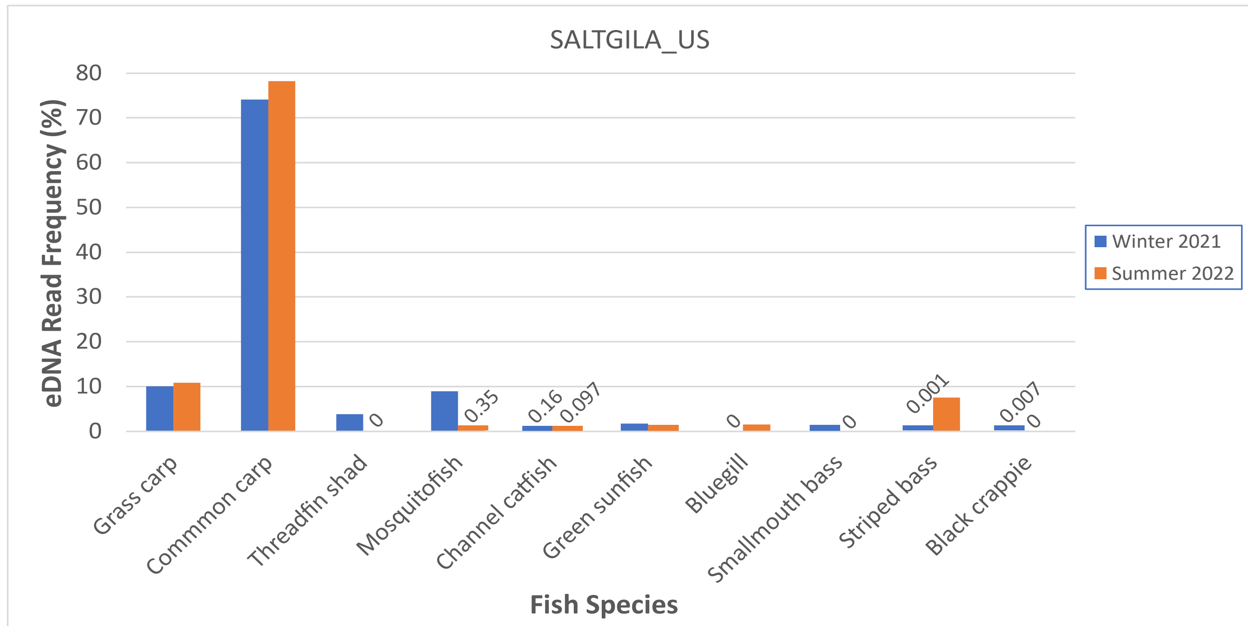

Figure 25: Percent read frequencies for fish species detected from eDNA metabarcoding from SALTGILA_US. Reads from samples collected in the winter of 2021 are displayed in blue. Reads from samples collected in the summer of 2022 are shown in orange. For species where the read frequency was below 1%, the value is displayed above bar.

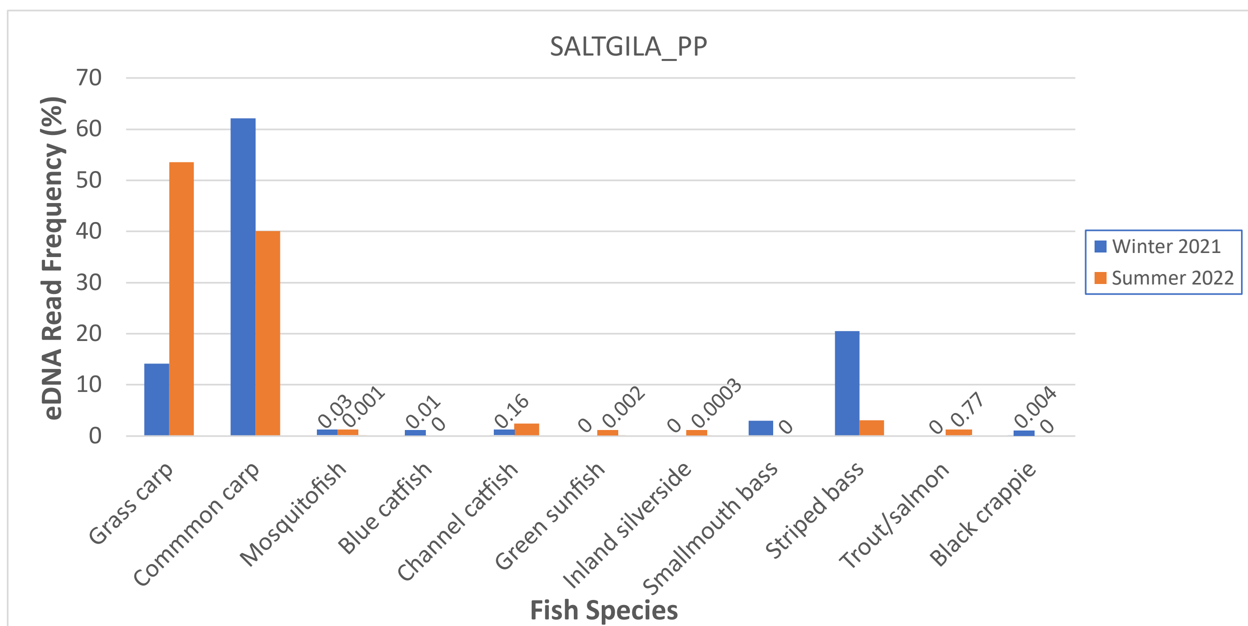

Figure 26: Percent read frequencies for fish species detected from eDNA metabarcoding from SALTGILA_PP. Reads from samples collected in the winter of 2021 are displayed in blue. Reads from samples collected in the summer of 2022 are shown in orange. For species where the read frequency was below 1%, the value is displayed above bar.

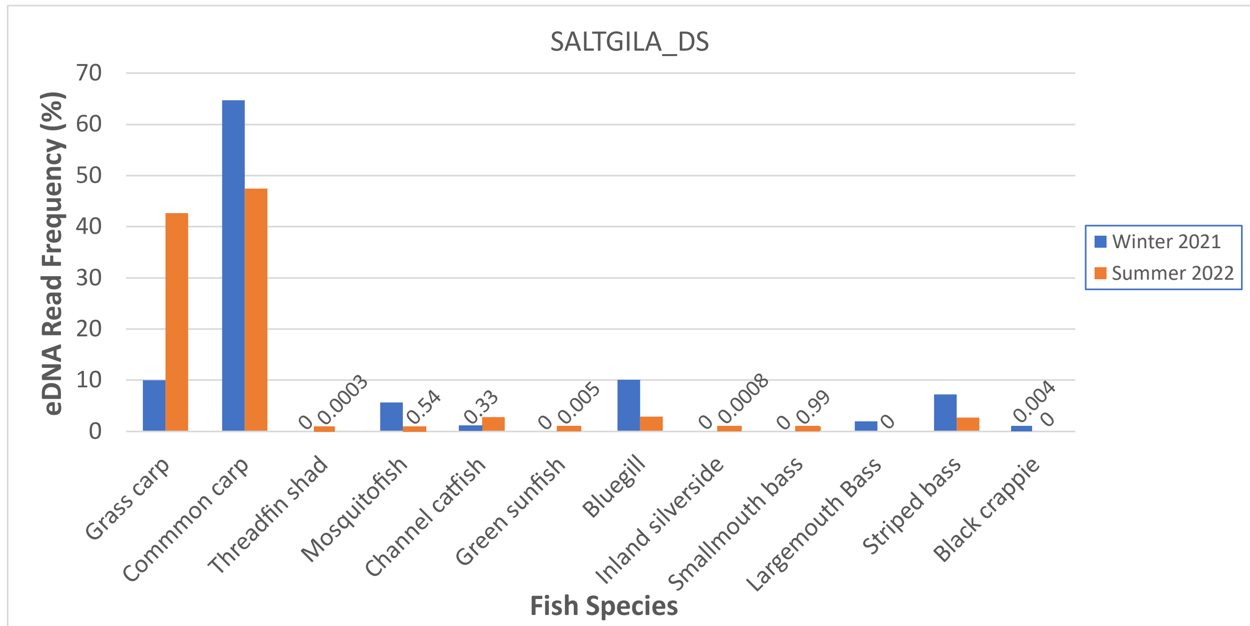

Figure 27: Percent read frequencies for fish species detected from eDNA metabarcoding from SALTGILA_DS. Reads from samples collected in the winter of 2021 are displayed in blue. Reads from samples collected in the summer of 2022 are shown in orange. For species where the read frequency was below 1%, the value is displayed above bar.

IV. References

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and S.P. Holmes. "DADA2: High-resolution sample inference from Illumina amplicon data." *Nature Methods*, 13 (2016), 581-583.

Carim, K. J., T. Wilcox, M. K. Young, K. S. McKelvey, and M. K. Schwartz. "Protocol for collecting eDNA samples from streams [Version 2.3]." *Boise, ID: US Department of Agriculture, Forest Service, Rocky Mountain Research Station, Boise Aquatic Sciences Lab. 10 p. Online: http://www. fs. fed. us/research/genomics-center/docs/edna/edna-protocol. pdf* (2015).

Clarkson, R.W., B.R. Kesner, and P.C. Marsh. "Long-Term Monitoring Plan for Fish Populations in Selected Waters of the Gila River Basin, Arizona. Revision 3." *U.S. Bureau of Reclamation, Phoenix, AZ* (2011).

Gill, C. and A. Jones. "Lake Pleasant Fish Management Plan." *Arizona Game and Fish Department, Phoenix, AZ*. (2019)

Martin, Marcel. "Cutadapt removes adapter sequences from high-throughput sequencing reads." *EMBnet. journal* 17, no. 1 (2011): 10-12.

Miya, Masaki, Y. Sato, T. Fukunaga, T. Sado, J. Y. Poulsen, K. Sato, Toshifumi Minamoto et al. "MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species." *Royal Society open science* 2, no. 7 (2015): 150088.

Mueller, G. "Establishment of a Fish Community in the Hayden-Rhodes and Salt-Gila Aqueducts, Arizona." *North American Journal of Fisheries Management,*16, no. 4 (1996): 795-804.

Passamaneck, Y. "Environmental DNA metabarcoding of fish in the Central Arizona Project canal for detection of non-native species." *Bureau of Reclamation, Denver, CO* (2022)

RStudio Team. RStudio: Integrated Development for R. *RStudio, PBC, Boston, MA.* http://www.rstudio.com/. (2020)

Stewart B.S., Meding, M.M., and D.R. Rogers. "Lake Pleasant striped bass." *Arizona Game and Fish Department, Research Branch, Technical Guidance Bulletin* No. 11. (2007)

USFWS (U.S. Fish and Wildlife Service). "Endangered species act section 7 Biological Opinion on transportation and delivery of Central Arizona Project water to the Gila River basin 2-21-90-F-119, April 15, 1994." *U.S. Fish and Wildlife Service, Phoenix, AZ* (1994).

USFWS (U.S. Fish and Wildlife Service). 2001. "Background information on the Central Arizona Project and nonnative aquatic species in the Gila River basin (excluding the Santa Cruz River subbasin)." *U.S. Fish and Wildlife Service, Phoenix, AZ* (2001).

USFWS (U.S. Fish and Wildlife Service). “Reinitiated biological opinion on transportation and delivery of Central Arizona Project water to the Gila River basin in Arizona and New Mexico and it potential to introduce and spread nonindigenous aquatic species.” *U.S. Fish and Wildlife Service, Phoenix, AZ* (2008).

Appendix 1

Table of sampling site information for eDNA metabarcoding survey

Site Name	Latitude	Longitude	Site Description	Reads §
HASS_US	33.64448	-112.711	CAP Canal, upstream of Waddell Canal	12,782,737
HASS_PP	33.66732	-112.691	Hassayampa Pumping Plant, upstream of Waddell Canal	11,657,842
HASS_DS	33.66834	-112.686	CAP Canal, upstream of Waddell Canal	10,413,993
HAWA_06	33.7798	-112.301	CAP Canal, upstream of Waddell Canal	10,140,896
WASG_01	33.78045	-112.252	Confluence of CAP Canal and Waddell Canal	10,333,842
WASG_02	33.77999	-112.25	CAP Canal, downstream of Waddell Canal	9,165,306
SALTGILA_US	33.52981	-111.705	CAP Canal, downstream of Waddell Canal	7,062,740
SALTGILA_PP	33.50519	-111.684	Salt Gila Pumping Plant, downstream of Waddell Canal	7,197,377
SALTGILA_DS	33.50031	-111.679	CAP Canal, downstream of Waddell Canal	8,058,073
LAKEPLEA_01	33.84889	-112.274	Lake Pleasant	9,662,868
WADD_PP	33.84168	-112.273	Waddell Pumping Plant	9,867,146
WADD_01	33.78366	-112.252	Waddell Canal	11,495,871

§ Total number of reads matched to a fish species