Environmental DNA Study in the East Bay, Spring 2021

Summary Report

Prepared **by R. Katznelson and M. Coyle**
With six Albany EDSET Interns
Final Report, November 23, 2021

Executive summary

"Environmental DNA" is a catch-all term for traces of DNA found in the environment. In creeks, environmental DNA is concentrated in sediments, arriving from many sources including intact microscopic organisms as well as hair, scales, eggs, feces, and other bodily sources. By sequencing these unique fragments of DNA, we can identify the organisms and therefore survey the biodiversity found in each location, with minimal disturbance to the organisms and their environment.

Beginning in February 2021, six Friends of 5 Creeks (F5C) interns from Albany High School's Environmental Design, Society, English, and Technology (EDSET) program for juniors and seniors worked with Dr. Revital Katznelson and PhD candidate Max Coyle on the F5C eDNA Project. The interns learned the essentials of experimental design, sampling, and DNA technology. They collected four mud samples from sediments of Cerrito Creek and Codornices Creek in the East Bay. The DNA fragments were extracted from the mud, purified, amplified to yield diagnostic DNA fragments, and sequenced, with the assistance of the CALeDNA project and Dr. Rachel Meyer's laboratory at UC Santa Cruz. The resulting sequences were assigned as best as possible to the closest specific taxonomic group (in some cases a particular species, in other cases a wider clade), allowing the identification of several hundred species. Additionally, in April 2021, the team collected multiple sediment samples from Codornices Creek to test for the presence of steelhead/rainbow trout (*Oncorhynchus mykiss)*. The DNA from these samples was extracted and tested for the presence of trout DNA using polymerase chain reaction (PCR) with trout-specific primers.

Our Codornices and Cerrito Creek samples revealed the presence of an impressive variety of microbes, fungi, algae, protozoa, crustaceans, insects, plants, and more. Some of these organisms are indicators of good water quality, and the richness of species is also a good sign. However, eDNA from Codornices Creek confirmed previous (Fitanides 2018) and current observations of an established population of the highly invasive New Zealand Mud Snail, and it also indicated the absence of a robust trout population. Among vertebrates (animals with backbones), human DNA, as well as DNA from rat and stickleback fish, was detected only in one mud sample that was successfully analyzed for vertebrates, but the DNA of these species was likely present in the other samples as well. Our diagnostic PCR test for the presence of trout was negative, consistent with the rarity of trout sightings in recent years. Given the limited resources for this project, our conclusions are limited to general observations.

1. Introduction

As animals and plants move through an environment, they leave behind traces of their presence in the form of feces, shed skin, feathers, fur, leaves, pollen, eggs, mucus, and more. All of these

contain “environmental DNA” (eDNA), which can be retrieved from samples such as water, soil, and sediment. DNA sequencing technology can then be used to identify the community of organisms present in that location. Over the last twenty years, with expanding databases of species-specific and group-specific DNA sequences, eDNA sequencing has begun development as a useful monitoring and research tool.

The CALeDNA Project (https://ucedna.com/) was set up to provide a comprehensive list of species found in multiple habitat types throughout California, covering different seasons across many years. As of September 2021, their database includes over 36,000 identified species from over 3,500 sites, and their data have been used for a number of biodiversity studies (https://ucedna.com/publications). Moreover, the project maintains a collection of frozen water, soil, and sediment samples, preserved for future studies that may track species’ responses to changes in environmental conditions over time. Sampling efforts have been extensively augmented by involving citizen scientists such as interns with the Friends of Five Creeks (Meyer *et al* 2021).

This report describes a short-term, exploratory project of citizen science inspired by the CALeDNA efforts. High-school students were trained to collect creek sediment samples which were used to characterize local flora and fauna based on presence of DNA, and thus augment the CALeDNA statewide database. The F5C interns also collected samples for analysis of trout DNA in Codornices Creek, two years after the spillover of a flame-fighting foam which killed hundreds of resident trout and left very few survivors.

The first study objective was characterization of biodiversity in urban creeks to augment statewide data, answering the question: **Which organisms shed detectable DNA in Cerrito Creek and Codornices Creek?**

The second objective was related to the aftermath of the 2019 fish kill in Codornices Creek, with the question: **Can trout DNA be found along lower Codornices Creek in the spring of 2021?**

2. Methods

2.1. Study design rationale and monitoring stations

Sampling locations were selected with specific study questions in mind, based on what we already know (i.e., using the “knowledge-based (directed) sampling design principle”.
To answer the **first question** – which organisms live in Cerrito Creek and Codornices Creek? – we tried to maximize detection of as many species as possible, so we selected creek segments with high habitat complexity, many shelter elements, and a variety of flow patterns (e.g., riffle and pools). For sediment collection we targeted depositional environments such as pools where we could find fine-grain sediments, because of reports that finer particles contain more concentrated eDNA (Barnes *et al* 2020), and from these samples we specifically enriched for the finest particulates (see below).

To answer the **second question** - can we find trout DNA in lower Codornices Creek? – we selected sampling stations in pools where trout had been observed before the fish kill event. We also collected samples in one location upstream of several migration barriers, where trout have

never been seen and are not expected to be now; this provided a "negative control" which allows the discrimination of potential false positives in our assay design (see below).

Table 2.1 shows the stations visited for sediment sampling in the spring of 2021.

Table 2.1: Monitoring stations used in the eDNA project, spring 2021

Waterbody	Station Name	Station ID	Latitude	Longitude	Station Location Description
Cerrito Creek	Cerrito Creekside Park at ford	203CER015	37.8982	-122.30563	200 m downstream of Creekside park, below weir
Cerrito Creek	Cerrito at Stannage	203CER025	37.89827	-122.29945	250 m Upstream of San Pablo Ave
Codornices Creek	Codornices below 5th Street	203COD029	37.88233	-122.30386	60 m downstream of footbridge at the end of 5th Street
Codornices Creek	Codornices below 6th Street	203COD032	37.88239	-122.30262	20 m downstream 6th St. bridge
Codornices Creek	Codornices at 6th Street	203COD033	37.88245	-122.30211	10 m east of 6th St bridge
Codornices Creek	Codornices at 8th Street	203COD038	37.88283	-122.30037	Downstream end of the 8th St. bridge
Codornices Creek	Codornices above 8th Street	203COD040	37.88268	-122.29896	At the end of 9th St near baseball diamond, about 150m upstream of 8th St.
Codornices Creek	Codornices above 10th Street	203COD046	37.88318	-122.29744	100 m Upstream of 10th St., a deep pool
Codornices Creek	Codornices at Live Oak Park	203COD119	37.88422	-122.26952	Live Oak Park downstream of Walnut St, under footbridge closest to Shattuck

2.2. Field operations

Field activities were conducted during winter base flow, usually a few days after a storm runoff event, during daylight (around mid-day). The time of year was constrained by 2021 EDSET internship schedule.

2.2.1 Field observations and measurements

Field observations using the CEDEN protocol were conducted in every site visit, augmented by measurements of water temperature and pH. Temperature was measured with a bulb thermometer (resolution 1 degree C and range between minus 20 and plus 50 degrees C), and pH was determined with pH strips made with 3 separate color pH indicators to encompass the range of pH 5 to pH 12 at 0.5 pH unit intervals.

2.2.2 Collection of sediment samples

A preliminary experiment was conducted early in the project (February 17th) to inform the selection of an appropriate sampling method for eDNA in sediments. After verification that fine sediments indeed hold more DNA, a procedure to collect only the particles < 220 microns (discarding sand and gravel) was adopted.

Table 2.2 shows the sampling log for the experiment and for the two subsequent collection events, on February 25 and April 22, 2021.

Table 2.2: Samples collected during the eDNA project, spring 2021

Station ID	Collection Date	Collection Time	Matrix	Sample type	Sample ID	Number of replicate	Fraction collected	sampling device	Number of Sub-samples
203COD038	17/Feb/2021	11:20	sediment	grab	D1,D2,D3	3	whole	spoon	1
203COD038	17/Feb/2021	11:30	sediment	composite	S1,S2,S3	3	<220 um	ladle	12
203CER015	25/Feb/2021	13:00	sediment	composite	CER015-1,2,3	3	<220 um	ladle	10
203CER025	25/Feb/2021	13:50	sediment	composite	CER025-1,2,3	3	<220 um	ladle	10
203COD032	25/Feb/2021	11:20	sediment	composite	COD032-1,2,3	3	<220 um	ladle	14
203COD038	25/Feb/2021	12:10	sediment	composite	COD038-1,2,3	3	<220 um	ladle	10
203COD029	22/Apr/2021	11:45	sediment	composite	COD029-1,2,3	3	<220 um	ladle	5
203COD033	22/Apr/2021	12:15	sediment	composite	COD033-1,2,3	3	<220 um	ladle	7
203COD040	22/Apr/2021	13:00	sediment	composite	COD040-1,2,3	3	<220 um	ladle	6
203COD046	22/Apr/2021	13:40	sediment	composite	COD046-1,2,3	3	<220 um	ladle	6
203COD119	22/Apr/2021	14:20	sediment	composite	COD119-1,2,3	3	<220 um	ladle	5

Sampling technique: A stainless steel ladle was used to scrape the top 1-2 cm of bed sediment under water, and the content was transferred to a 220-micron sieve lodged inside a glass bowl. The sieve was shaken inside the bowl with very little water, and then the sieve's content (sands and gravels) was discarded. A fresh aliquot – another ladle – of surface sediment was added to the sieve and shaken in the liquid, and then the large particles were discarded. Each sample was comprised of 5 to 14 such aliquots, depending on the sediment properties. The resulting composite sample was mixed thoroughly and continuously while small increments were dispensed into vials for DNA extraction, and the remaining slurry was used to determine percent moisture.

2.3. Laboratory Analyses

To answer the first question (who lives in our two urban creeks) we needed purified eDNA for analysis using a high-throughput shotgun sequencing approach. For this purpose, environmental DNA was extracted, purified, amplified, and sequenced by UC Santa Cruz personnel (as part of the CALeDNA project). The second study question (presence or absence of trout DNA) could be answered by targeted amplification of species-specific sequences.

2.3.1 Extraction and purification of DNA

Question 1: DNA extraction for high-throughput shotgun sequencing:

This was performed in the CAL eDNA laboratory at UC Santa Cruz. The protocol uses the Qiagen DNeasy Powersoil Kit and is adapted from the published protocol for that kit. Detailed information about DNA extraction for high-throughput sequencing can be found here: https://ucedna.com/methods-for-researchers

Extraction and purification procedures involved these 4 steps:
Step 1: Pieces of DNA were extracted from sediment samples by three forces: (a) vigorous shaking with tiny steel balls (about 0.5 mm diameter) to break sediment aggregates and pulverize cell structures, including cell walls; (b) detergents that dissolve membranes, breaking apart cells and sub-cellular organelles; and (c) salts that disrupt electrostatic forces that bind DNA to particulate matter. The steel balls, sediment particles, and cell debris was then removed by centrifugation.
Step 2: Now that the (water-soluble) DNA pieces were in aqueous solution, non-DNA molecules could be removed by selective precipitation, followed by centrifugation to remove the precipitates.
Step 3: When ethanol was added in the next step, the DNA pieces became insoluble and were allowed to attach to a solid matrix (a filter membrane); other remaining molecules that are soluble in ethanol could then be washed away.
Step 4: The purified DNA pieces were then eluted from the membrane by dissolving them in water again.

Question 2: DNA extraction for targeted amplification of steelhead-specific eDNA markers.

DNA was extracted using a CTAB and chloroform:isoamyl protocol as described in Turner et al, 2015. Briefly, sediment samples were suspended in CTAB buffer, vortexed, and heated. Chloroform:isoamyl alcohol mixture was used to pull out non-aqueous impurities. DNA was precipitated from the aqueous fraction with isopropyl alcohol and salt, then washed with 70% ethanol, and resuspended in water.

For both DNA extraction methods (Qiagen Powersoil and CTAB), more detailed protocols are available upon request.

2.3.2 Hybridization with DNA Primers for the amplification of selected DNA fragments

Both high-throughput shotgun sequencing (Question 1) and detection of trout-specific DNA (Question 2) required amplification of specific genomic regions. These regions are selectively amplified by short DNA primers that hybridize on either end of these genomic regions. Thus, the genomic DNA between two primer binding sites is amplified, making it accessible to further analysis. Some primers target and amplify DNA regions that might be shared by a group of organisms, e.g., all vertebrates (Question 1), and later it is necessary to sequence each amplified fragment to determine the DNA sequence differences that will tell apart one taxon from another. Other primers are specific to certain organisms (e.g., trout, Question 2) and therefore PCR amplification should only happen if that specific piece of DNA is present in the sample. In this way, the presence or absence of a PCR product is used as a proxy to assess the presence of a specific genomic DNA fragment in the sample.

We used the following generic primer sets for metabarcoding (high-throughput shotgun sequencing, Question 1); each amplifies genomic regions from the associated groups listed next to the Primer names:

- 16S = bacteria and archaea
- 18S = general eukaryotes
- PITS = plants and green algae
- FITS = fungi

- CO1 = brown algae, other protists, and invertebrates
- Vert12S = vertebrates

Exact sequences can be found here: https://ucedna.com/methods-for-researchers

For Question 2, the primers used for detection of steelhead trout eDNA (Question 2) were intended to be specific to steelhead genomes, and therefore the presence of a PCR product indicated successful DNA amplification, i.e., a match between the primer sequences and DNA present in the environmental sample. Successful amplification was determined by running the PCR product on a gel electrophoresis platform: if it formed a clear band, we could conclude that we had specificity for the target sequence, and finding the band at the correct location on the gel indicated that the fragment had the correct size.

For detection of steelhead eDNA, five sets of previously published primers were used. These are listed below, with the genomic region that is targeted (e.g. COI) as well as the last name of the first author and DOI for the associated publication. Note that because of time and resource limitations, regular PCR (and not quantitative PCR with fluorescent probes) was used for detection. This could lead to false positive results, but likely not false negatives.

CCTCCCGTGAGGACAAATATCA	O. mykiss cyt B F, Hernandez et al 2020, 10.1002/edn3.89
TGGCGTTGTCAACGGAGAAG	O. mykiss cyt B R, Hernandez et al 2020, 10.1002/edn3.89
AACATAAAACCTCCAGCCATCTCT	O. mykiss COI F, Brandl et al 2014, 10.1111/1755-0998.12305
AGCACGGCTCAAACGAAAA	O. mykiss COI R, Brandl et al 2014, 10.1111/1755-0998.12305
CTTTCTCCTCCTCCTGTCTTCA	O. mykiss COI F, Wood et al 2020, 10.1002/edn3.64
GAAATTCCAGCTAAATGAAGGGAG	O. mykiss COI R, Wood et al 2020, 10.1002/edn3.64
AGCCACCTCATTTACTGCCATT	O. mykiss NAD5 F, Wood et al 2020, 10.1002/edn3.64
CGGGGGTGTTGGTGGGTAGGAA	O. mykiss NAD5 R, Wood et al 2020, 10.1002/edn3.64
AGTCTCTCCCTGTATATCGTC	O. mykiss NADH region F, Wilcox et al 2015, 10.1371/journal.pone.0142008
GATTTAGTTCATGAAGTTGCGTGAGTA	O. mykiss NADH region R, Wilcox et al 2015, 10.1371/journal.pone.0142008

Each test requires that we also run a negative control, i.e., a sample we are sure does not contain the DNA sequence of trout. For that we used sediment samples from a Codornices Creek location that was 1) upstream from any previous steelhead sighting and 2) almost certainly unreachable by migrating trout due to creek blockages. We also ran a positive control by extracting DNA from trout tissue purchased from a local grocery store.

2.3.4 Sequencing of individual DNA fragments.

For metabarcoding, PCR products had Nextera indices added to them and were sequenced with Illumina technology. More information here: https://ucedna.com/methods-for-researchers

2.4 Data Quality Checks

To collect and share data of known quality, researchers must perform a variety of actions to check, record, and report the quality of the data. These actions were different for different methods in our study, and included the following:

Field measurements: conduct comparison with standards to assess accuracy; repeat measurements of same to assess precision.
Steelhead/trout eDNA amplification: Positive control: steelhead/trout tissue from a grocery store. Negative control: samples from Codornices Creek location above the dams that preclude upstream migration.
Sequencing results: If samples are missing, it means they did not produce enough reads to indicate acceptable quality, and therefore were not assigned to taxonomic groups and are not part of our data set. In our study we found that plants and vertebrates were especially difficult to detect, possibly due to poor amplification with the primer sets for these groups.

3. Results and Discussion

Though preliminary in nature, this project generated satisfactory answers to the two study questions and highlighted the utility of eDNA sampling. The highlights of eDNA findings are described below, while all other data (including measurements and observations) are available upon request.

3.1 Biodiversity and Species of Interest in Cerrito and Codornices Creeks.

The results from CALeDNA sequencing of our sediment samples suggested almost 1000 unique species identified across our sampling sites in Codornices and Cerrito Creeks. Table 3.1 shows some examples taken from the Project's data spreadsheet, and the entire data set can be found here: http://fivecreeks.org/background/eDNA_spreadsheet.pdf

Table 3.1: Selected examples of orgnisms whose DNA was detected in Cerrito and Codornices Creeks

Phylum	Class	Order	Family	Genus	Species	what it is	Metabarcode
Annelida	Clitellata	Haplotaxida	Naididae	Tubifex	Tubifex tubifex	worm	18S-general eukaryotes
Apicomplexa	Conoidasida	Eucoccidiorida	Cryptosporidiidae	Cryptosporidium	Cryptosporidium parvum	pathogen	18S-general eukaryotes
Arthropoda	Hexanauplia	Calanoida	Temoridae	Eurytemora	Eurytemora affinis	copepod	18S-general eukaryotes
Arthropoda	Hexanauplia	Harpacticoida	Thalestridae	Eudactylopus	Eudactylopus sp.	copepod	18S-general eukaryotes
Arthropoda	Insecta	Coleoptera	Scarabaeidae	Perissosoma	Perissosoma sp.	beetle	18S-general eukaryotes
Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium digitatum	mold	FITS-Fungi
Ascomycota	Sordariomycetes	Hypocreales	Nectriaceae	Fusarium	Fusarium oxysporum	mold	CO1-protists-invertebrates
Bacillariophyta	Bacillariophyceae	Bacillariales	Bacillariaceae	Nitzschia	Nitzschia amphibia	diatom	18S-general eukaryotes
Bacillariophyta	Bacillariophyceae	Cymbellales	Gomphonematace	Gomphonema	Gomphonema parvulum	diatom	CO1-protists-invertebrates
Bacillariophyta	Bacillariophyceae	Naviculales	Naviculaceae	Navicula	Navicula ramosissima	diatom	16S-bacteria and archaea
Bacillariophyta	Bacillariophyceae	Naviculales	Pinnulariaceae	Pinnularia	Pinnularia sp.	diatom	CO1-protists-invertebrates
Bacillariophyta	Bacillariophyceae	Thalassiophysales	Catenulaceae	Amphora	Amphora normanii	diatom	18S-general eukaryotes
Bacillariophyta	Coscinodiscophyc	Melosirales	Melosiraceae	Melosira	Melosira varians	diatom, filamentous	16S-bacteria and archaea
Chordata	Mammalia	Primates	Hominidae	Homo	Homo sapiens	human	vert12S-vertebrates
Chordata	Mammalia	Rodentia	Muridae	Rattus	Rattus norvegicus	rat	vert12S-vertebrates
Ciliophora	Heterotrichea	Heterotrichida	Stentoridae	Stentor	Stentor roeselii	ciliate	18S-general eukaryotes
Cnidaria	Hydrozoa	Limnomedusae	Olindiidae	Craspedacusta	Craspedacusta sowerbii	jellyfish	18S-general eukaryotes
Mollusca	Gastropoda	Littorinimorpha	Tateidae	Potamopyrgus	Potamopyrgus antipodarum	New Zealand mud snail	CO1-protists-invertebrates
Nematoda	Chromadorea	Rhabditida	Spirocercidae	Mastophorus	Mastophorus muris	parasite-of rodents	18S-general eukaryotes
Platyhelminthes	Rhabditophora	unknown	Bothrioplanidae	Bothrioplana	Bothrioplana sinensis	leech	18S-general eukaryotes
Rotifera	Bdelloidea	Philodinida	Philodinidae	Philodina	Philodina megalotrocha	rotifer	18S-general eukaryotes
Streptophyta	unknown	Cupressales	Cupressaceae	Juniperus	Juniperus chinensis	juniper tree	PITS-plants and green alga
Tardigrada	Eutardigrada	Parachela	Hypsibiidae	Hypsibius	Hypsibius convergens	water-bear	18S-general eukaryotes

Whereas previous monitoring activities had provided ample information on insects and other animals found in our creeks, the results from this study represent the first thorough accounting of microscopic life present in these two urban creeks: a variety of amoebae, diatoms, ciliates, and microscopic animals such as rotifers and tardigrades were found.
In addition to quantifying the number of distinct species present in these creeks, we also looked for taxa that would be particularly interesting or relevant with respect to the ecology of these environments:

- New Zealand Mud Snail *(Potamopyrgus antipodarume)* eDNA was detected in Codornices Creek, corroborating the macroscopic identification of these highly invasive gastropods.
- Another interesting species detected was the peach blossom jellyfish (*Craspedacusta sowerbii*) in Cerrito Creek. This is one of the only known freshwater jellyfish in the world, originating from China, but spreading all around the world in the last century.
- Additionally, we detected a nematode worm that parasitizes rodents, *Mastophorus muris* in Codornices Creek. We also detected its likely host, the common rat *Rattus norvegicus*. This is a nice example of how eDNA results can identify ecological interactions like host-parasite ones.
- A diversity of copepods were also found in our samples. These small crustaceans form an integral part of the food web in creek ecosystems.

Profile pages (with pictures and information) of these and other species were made for the purpose of communicating highlights of these eDNA results to broader audiences. Those profiles can be viewed here: http://fivecreeks.org/background/eDNA%20profiles.pdf

3.2 Steelhead/trout eDNA could not be detected by PCR amplification

Primer sets 3 and 4 gave non-specific PCR products, as evidenced by seeing bands for all five Codornices Creek sites, including COD120, our upstream negative control site (not shown). While it is theoretically possible that trout DNA was found at our negative control site (introduced by human activity or bird feces), non-specificity of the primer set is more plausible to us.

On the other hands, primer sets 1, 2, and 5 yielded no PCR product for any of the five Codornices Creek sites. To test whether this could be a false negative, we added exogeneous trout DNA (using tissue from a grocery store) to Codornices Creek sediment samples and carried them through the same DNA extraction and PCR amplification process. The results are shown in Figure 3.1, as bands on the gel electrophoresis platform. Primer sets 1 and 5 showed successful amplification when trout tissue was added to the sample, but no amplification in the un-modified sediment samples. The most plausible interpretation is that trout eDNA is not present in sufficient quantities in our sediment samples for PCR detection. It could be missing altogether or exceedingly sparse.

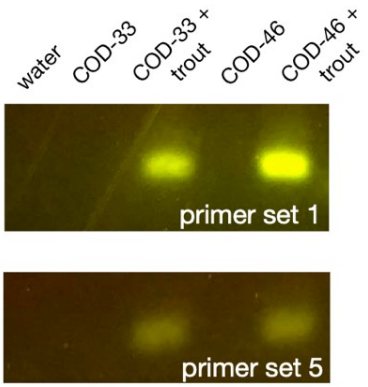

Figure 3.1 Gel electrophoresis bands of PCR products with and without spiked trout DNA

4. References

Barnes MA, Chadderton WL, Jerde CL, Mahon AR, Turner CR, Lodge DM (2020). **Environmental conditions influence eDNA particle size distribution in aquatic systems.** https://doi.org/10.1002/edn3.160

Fitanides 2018 The Watershed Project Benthic Macroinvertebrates Study (Results on CEDEN)

Meyer, R. S, Ramos, M. M, Lin, M., Schweizer, T. M, Gold, Z., Ramos, D. R, et al. (2021). **The CALeDNA program: Citizen scientists and researchers inventory California's biodiversity**. California Agriculture, 75(1). http://dx.doi.org/10.3733/ca.2021a0001
Retrieved from https://escholarship.org/uc/item/5qk7q93p

Turner, Cameron R., Karen L. Uy, and Robert C. Everhart. 2015. **"Fish Environmental DNA Is More Concentrated in Aquatic Sediments than Surface Water."** *Biological Conservation* 183 (March): 93–102.