

# **2018 GENETIC BARCODING AND ASSESSMENT OF ENVIRONMENTAL DNA FOR BIODIVERSITY SAMPLING WITHIN THE PORTS OF LOS ANGELES AND LONG BEACH**

**Prepared for:**

**Port of Los Angeles  
Environmental Management Division  
425 South Palos Verdes Street  
San Pedro, CA 90731**

**Port of Long Beach  
Environmental Planning Division  
415 W. Ocean Blvd.  
Long Beach, CA 90802**

**June 2020**

**Prepared by:**

Zachary Gold<sup>1</sup>, Kevin Stolzenbach<sup>2</sup>, Teia M. Schweizer<sup>3</sup>, Regina Wetzer<sup>4</sup>, N. Dean Pentcheff<sup>4</sup>, Adam R. Wall<sup>4</sup>, Emily E. Curd<sup>1</sup>, Rachel S. Meyer<sup>1</sup>, Robert Wayne<sup>1</sup>

<sup>1</sup>Department of Ecology and Evolutionary Biology, University of California Los Angeles, Los Angeles, CA 90095

<sup>2</sup>Wood Environment and Infrastructure, Inc. 9210 Sky Park Court, Suite 200, San Diego, CA 92123

<sup>3</sup>Department of Biology, Colorado State University, Fort Collins, CO 80521

<sup>4</sup>Diversity Initiative for the Southern California Ocean (DISCO), Natural History Museum of Los Angeles County, Los Angeles, CA 90007

---

## TABLE OF CONTENTS

<b>LIST OF TABLES .....</b>	<b>2</b>
<b>LIST OF FIGURES.....</b>	<b>3</b>
<b>LIST OF APPENDICES.....</b>	<b>4</b>
<b>ACRONYMS AND ABBREVIATIONS.....</b>	<b>5</b>
<b>1 INTRODUCTION.....</b>	<b>6</b>
1.1 Genetic barcoding and environmental DNA .....	6
1.2 Genetic barcode reference library generation .....	7
1.3 Spatial and temporal resolution of environmental DNA .....	8
1.4 Estimation of abundance with environmental DNA.....	8
1.5 Project objectives.....	9
<b>2 MATERIALS AND METHODS .....</b>	<b>10</b>
2.1 Overview of environmental DNA sampling and processing.....	10
2.2 Seawater sampling of environmental DNA.....	10
2.3 Trawl and other 2018 Biosurvey sampling.....	11
2.4 Contamination precautions .....	12
2.5 Environmental DNA extractions, amplification, and sequencing .....	12
2.6 Bioinformatic analysis and taxonomic determination.....	13
2.7 Bioinformatic decontamination .....	13
2.8 Normalizing environmental DNA reads for abundance estimation.....	14
2.9 Comparisons of taxon occurrences.....	14
2.10 Generation of barcode reference sequences .....	15
<b>3 RESULTS .....</b>	<b>16</b>
3.1 Seawater sampling of environmental DNA.....	16
3.2 Sequencing.....	17
3.3 Reference barcode generation.....	17
3.4 Survey comparisons for fish .....	18
3.5 Survey comparisons for invertebrates and algae .....	24
3.6 Detection of known invasive species .....	25
3.7 Detection of spatial variation by environmental DNA .....	27
3.8 Estimation of fish abundance and biomass from environmental DNA.....	29
<b>4 DISCUSSION.....</b>	<b>31</b>
4.1 Comparisons between environmental DNA and conventional sampling .....	31
4.1.1 Incomplete barcode reference libraries.....	31
4.1.2 Primer design and bioinformatic limitations .....	33
4.1.3 High sensitivity of eDNA approaches .....	33
4.1.4 Sampling biases, scope, and scale .....	35
4.2 Abundance and biomass .....	35
4.3 Environmental DNA identifies spatial variation .....	36
<b>5 PROSPECTS.....</b>	<b>38</b>
<b>6 REFERENCES CITED .....</b>	<b>41</b>
<b>APPENDIX I. INVERTEBRATE REFERENCE BARCODES .....</b>	<b>48</b>
<b>APPENDIX II: FULL METHODS .....</b>	<b>52</b>
Contamination precautions .....	52
Environmental DNA extractions, amplification and sequencing.....	52
Bioinformatic analysis and taxonomic determination .....	53
Bioinformatic Decontamination .....	54
Normalizing eDNA reads across samples .....	55
Barcode reference sequences generated for this project.....	55
R analysis methods — Statistical methods .....	56
<b>APPENDIX III: MAPS OF SAMPLING STATIONS AND LOCATIONS .....</b>	<b>57</b>

---

## LIST OF TABLES

Table 2-1. Sampling design for eDNA seawater sampling.....	11
Table 3-1. Seawater samples taken at each location at each station for eDNA analysis.....	16
Table 3-2. Summary of sequences generated from eDNA metabarcoding. ....	17
Table 3-3. Fish barcode references created for this project.....	17
Table 3-4. Enumeration of fish taxa observed in parallel trawl and eDNA samples. ....	19
Table 3-5. Fish taxa observed in both eDNA data and Biosurvey. ....	21
Table 3-6. Fish taxa observed only in Biosurvey data. ....	21
Table 3-7. Fish taxa observed only in eDNA data deemed reasonable as correct matches. ....	22
Table 3-8. Additional fish taxa observed only in eDNA data.....	23
Table 3-9. Invasive species from the Cal-NEMO Database observed in eDNA data Biosurvey. ....	26

---

## LIST OF FIGURES

Figure 2-1. Flow chart of eDNA sample processing.....	10
Figure 3-1. Counts of fish taxa observed in the eDNA data and trawl sampling. ....	20
Figure 3-2. Counts of fish taxa observed in the eDNA data and Biosurvey.....	20
Figure 3-3. Species rarefaction curve for eDNA sampling of fish species.....	24
Figure 3-4. Counts of invertebrate taxa observed in the eDNA data and Biosurvey. ....	24
Figure 3-5. Counts of invertebrate genera observed in eDNA data and Biosurvey.....	24
Figure 3-6. Counts of invertebrate families observed in eDNA data and Biosurvey.....	24
Figure 3-7. Rarefaction curve for eDNA sampling of all invertebrate taxa. ....	25
Figure 3-8. Counts of known invasive species observed in eDNA and Biosurvey. ....	25
Figure 3-9. Count of fish species observed in eDNA data at each station.....	27
Figure 3-10. Count of all species observed in eDNA data at each station. ....	27
Figure 3-11. Dissimilarity across biological replicates and sites.....	28
Figure 3-12. Bray-Curtis dissimilarity ordination for all taxa.....	28
Figure 3-13. Apportioned variance for eDNA samples.....	29
Figure 3-14. eDNA index compared to trawl fish abundance. ....	30
Figure 3-15. eDNA index compared to trawl fish biomass. ....	30

---

## **LIST OF APPENDICES**

Appendix I. Species list of marine invertebrate barcode voucher specimens generated for this project.

Appendix II. Full methods.

Appendix III. Maps of sampling stations and locations.

---

## ACRONYMS AND ABBREVIATIONS

<b>ASV</b>	Amplicon sequence variant
<b>BOLD</b>	International Barcode of Life Data System, a sequence database specifically devoted to DNA barcoding
<b>Bioinformatics</b>	Computational processing of DNA sequences. In this specific context, refer to it as taking raw DNA sequences and processing them to species level community data
<b>COI</b>	The cytochrome oxidase subunit 1 gene
<b>bp</b>	One DNA base pair
<b>Decontamination</b>	The process of removing DNA sequences which were unintentionally detected in negative controls and field blanks through bioinformatics
<b>DISCO</b>	Diversity Initiative for the Southern California Ocean, an eDNA and genetic barcoding program of the Natural History Museum of Los Angeles County
<b>DNA Barcoding</b>	Identification of species using a standardized fragment of DNA
<b>eDNA</b>	Environmental DNA
<b>GenBank</b>	Open access sequence database maintained by the National Center for Biotechnology Information
<b>Index Hopping</b>	A phenomenon on many next generation sequencing platforms where a sequence from one sample can “hop” to any other sample that is being sequenced during the same sequencing run on a sequencing machine. This “hopping” can lead to a sequence being miss-assigned to a sample it was not from. This phenomenon can be controlled with proper experimental design.
<b>Primer</b>	A short DNA fragment that targets a specific region of DNA for DNA synthesis and amplification
<b>PCR</b>	Polymerase chain reaction, used to amplify the number of copies of a selected gene in a sample
<b>Metabarcoding</b>	DNA barcoding of multiple species from the same environmental sample, usually using established DNA barcoding primers
<b>NHMLA</b>	Natural History Museum of Los Angeles County

---

## **1 INTRODUCTION**

The Ports of Los Angeles and Long Beach have performed extensive surveys of biodiversity in the Ports in 2000, 2008, 2013–14, and most recently in 2018. These surveys have been highly informative regarding the quality of biological habitat within the Port Complex. Conventional marine biodiversity surveys depend on physical sampling of organisms, whether by hand using divers, or by various trawl, seine, or grab methods. Following collection, specimens must be identified by expert taxonomists, either on board or later on shore following preservation, which can be time and cost intensive.

### **1.1 Genetic barcoding and environmental DNA**

Modern genetic techniques provide a new approach to complement conventional surveying. Genetic “barcoding”, developed over the last two decades, uses the diversity in a handful of genes shared across all organisms to distinguish between taxonomic groups (Taberlet et al. 2018). Genes have been identified that have both conserved regions that are shared across multiple species as well as variable regions that differ between species. Conserved regions allow for the design of molecular primers that permit genes of interest to be recognized and isolated across broad taxonomic groups; the variable regions permit differentiation between species. Therefore these “barcoding” gene regions can be targeted and sequenced, identifying specimens to species or a higher taxonomic level. Different genes are used for plants and algae, and animal taxa are distinguished using genes depending on the genetic and evolutionary similarities between species. For most invertebrates, the common barcoding genes are a fragment of the mitochondrial cytochrome oxidase subunit 1 gene (COI) and the 16S ribosomal DNA (16S). To distinguish closely related fish species, a mitochondrial fragment of the 12S ribosomal DNA (12S) is used.

Importantly, being able to use a sequence to identify a sequence to the species level depends on: 1) having a reference sequence for that species available in a searchable database, and 2) having sufficient differences in the DNA sequence between species to allow for identification. Thus the creation of reference sequences for species is an important undertaking to improve the taxonomic resolution of DNA barcoding techniques. Reference sequences must be derived from specimens that are collected, reputably determined to the species level by experienced taxonomists, then vouchered, sequenced, and registered in a widely accessible database (Taberlet et al. 2018). This process depends on time and cost intensive conventional sampling techniques.

As sequencing costs have dramatically declined in the past decade, novel applications of DNA barcoding have developed (Deiner et al. 2017). Using high throughput sequencing, it is possible to simultaneously sequence all of the DNA barcode sequences present in a sample, yielding hundreds or thousands of unique sequences from a single sample, a process known as *metabarcoding* (Taberlet et al. 2012b). A growing application of metabarcoding is “environmental DNA” (eDNA) which relies on the fact that all organisms shed DNA into the environment (in the form of cells or as free DNA) allowing for a sample of the environment (e.g. seawater) to be collected, filtered, extracted, amplified, and sequenced to determine which organisms have been in the area recently

(Deiner et al. 2017). eDNA metabarcoding cross-references sequences collected from the environment against reference databases of sequences to generate an inventory of taxon names (Bohmann et al. 2014). Reference databases are generally assembled from one or more of GenBank and BOLD public sequence repositories (Ratnasingham and Hebert 2007), and custom-assembled databases from vouchered and barcoded specimens (Machida et al. 2017). Where species-level reference barcodes exist, and species have distinct barcodes, identification to the species level can be made. Where no reference barcode for a species exists or where barcodes between two closely related species are identical, it is usually possible to at least give a higher-order taxonomic identification (to the level of genus, family, or above) (Curd et al. 2019). Advances in eDNA techniques have allowed scientists to accurately detect over 93% of known fish species from seawater in both aquaria and reef ecosystems (Kelly et al. 2014, Miya et al. 2015). In addition, previous research has found that eDNA metabarcoding approaches are highly sensitive to the detection of marine species, capturing broader diversity of native and locally relevant marine species than traditional fish survey methods, including observer transect surveys and trawling (Thomsen et al. 2012a, Port et al. 2016, Valentini et al. 2016).

In addition, eDNA methods have some clear advantages over traditional marine biodiversity survey methods. First, the approach uses small volumes of seawater and standard microbiology filtering techniques, allowing for simple and rapid collection of samples even in remote locations (Valentini et al. 2016, Deiner et al. 2017). Second, eDNA surveys have no measurable impact on marine ecosystems because it is a “takeless” sampling protocol (after the initial reference barcodes are registered) (Beja-Pereira et al. 2009). A conventional trawl may sacrifice thousands of fish, whereas an eDNA sample is a single bottle of water. Third, seawater eDNA metabarcoding can be extremely effective in detecting rare and cryptic species that are frequently missed with traditional fish biodiversity assessment methods (Port et al. 2016, Valentini et al. 2016, Boussarie et al. 2018), including both endangered species and recent introductions of invasive species, which are of important conservation management concern (Taberlet et al. 2012a, Bohmann et al. 2014, Kelly et al. 2014, Gold 2020).

## 1.2 Genetic barcode reference library generation

There is tremendous global investment to get DNA sequences for all families, genera and species that would enable highly sensitive and accurate molecular biomonitoring (e.g. Earth Biogenome Project) (Lewin et al. 2018). While reference databases are still incomplete, informatic analyses leverage similar sequences to assign sequences to genus, family, or higher classification levels, which still allows taxonomic biodiversity to be counted even if it cannot all be named. The better the coverage in the database for the taxa in the sampling area, the more precise the eDNA inventory will be. Hence significant effort must be invested in building regionally complete reference databases.

Because of variability within and between species and populations, it is necessary to barcode multiple individuals of a species from known localities to validate that we are amplifying and sequenc-

ing the species that we are calling by a single Linnean name, and to allow for discovery of undescribed, unnamed species that have not been previously recognized. Ideally having 4–5 high quality barcode sequences for each species is desirable.

Comparison of sequences from eDNA surveys against barcode reference databases presents a high-throughput opportunity to find leads for undescribed species — species new to science and not yet recognized as distinct (Appeltans et al. 2012). Unrecognized and cryptic species are estimated to be more prevalent in lesser-studied groups, a problem prevalent across the invertebrates (Beheregaray and Caccone 2007, Pfenninger and Schwenk 2007, Karanovic et al. 2016, Pérez-Ponce de León and Poulin 2016).

### **1.3 Spatial and temporal resolution of environmental DNA**

Preliminary studies in marine and freshwater settings indicate the degradation of eDNA occurs in hours to a few days (Dejean et al. 2011, Thomsen et al. 2012a, 2012b, Barnes et al. 2014, Barnes and Turner 2016, Sassoubre et al. 2016, Collins et al. 2018, Murakami et al. 2019). This suggests that eDNA samples capture a snapshot of marine diversity over a relatively small integration time, only sampling species that recently inhabited or passed through an area. This is consistent with research in Monterey Bay which has highlighted the ability of eDNA to accurately distinguish communities at the microhabitat level (<60 m apart), suggesting that biological not physical controls are the dominant driver of eDNA dynamics in kelp forest ecosystems (Port et al. 2016). Interestingly, the Monterey Bay study was conducted in a semi-enclosed, low velocity marine environment, similar to the Ports of Los Angeles and Long Beach, where advection and transport of eDNA was likely low (Port et al. 2016). In contrast, more recent work conducted in Maizuru Bay, Japan and Puget Sound, WA, both more dynamic marine systems, found spatial variation on the scales of 500–1,000 m (O'Donnell et al. 2017, Yamamoto et al. 2017). Importantly, in the Maizuru Bay study, previous researchers identified strong signal from a local fish processing plant with the transport of fish eDNA extending up 500 m away (Yamamoto et al. 2017), an important consideration for the Ports of Los Angeles and Long Beach which support a large seafood industry and have two large aquarium facilities in close proximity (Aquarium of the Pacific in Long Beach and Cabrillo Marine Aquarium in San Pedro). Together these results strongly suggest an important role for both biological and physical processes in governing the fate and transport of eDNA in marine systems with more exposed and dynamic marine environments likely experiencing greater transport and mixing of eDNA signatures.

### **1.4 Estimation of abundance with environmental DNA**

There remains uncertainty in the ability of eDNA to accurately measure relative abundance (Deiner et al. 2017). Recent studies in freshwater ecosystems have found species-specific eDNA sloughing rates and eDNA persistence, suggesting the need for species-specific relative abundance equations, potentially complicating the use of eDNA for estimating abundance (Evans et al. 2016, Sassoubre et al. 2016). However, despite these limitations, there is some evidence for the success

of this technique in estimating relative abundance or biomass under well-understood circumstances (Kelly 2016, Port et al. 2016, Thomsen et al. 2016, Taberlet et al. 2018, Tillotson et al. 2018). Recent theoretical and empirical research has found that the largest factor influencing accurate relative abundance estimates is PCR bias during the amplification process of eDNA (Kelly et al. 2019). PCR amplification bias is driven by the chemistry of DNA primer binding, with different species having different affinities for a given DNA primer. The process of exponential amplification of DNA sequences during PCR compounds this bias exponentially, explaining the vast majority of variation in the relative abundance of DNA reads. By making a reasonable assumption from our knowledge of molecular biology, namely that primer bias is controlled by chemistry and consistent across eDNA samples, we can use data transformation techniques to generate an “eDNA index” to allow us to compare a species-specific index of abundance across all given samples (Kelly et al. 2019). That is, though abundance comparisons between species are still uncertain, it may be possible to compare abundances within a species between sampling sites or times. Using this eDNA index method has been shown to dramatically improve eDNA abundance estimates, in some cases showing that eDNA methods have lower variance in abundance estimates than traditional beach seine methods (Tillotson et al. 2018, Kelly et al. 2019). These recent advances strongly suggest that strategically deployed eDNA approaches have the potential to provide complementary and independent relative abundance estimates of marine species.

## 1.5 Project objectives

The aim of this study was to directly compare eDNA metabarcoding approaches to concurrently conducted conventional benthic trawl sampling to identify the advantages and limitations of both methods. In addition to direct trawl comparisons, we compared biodiversity results of eDNA samples to those collected from all 2018 Biosurvey data. Together these comparisons will provide an examination of the extent of consistency or complementarity between survey methods.

In addition to eDNA surveys, specimens derived from the benthic sampling program (and from other sources) collected during the 2018 Biosurvey program were processed by the DISCO program to create new genetic barcode reference sequences to improve eDNA methods and allow for a more robust and accurate comparison of methods.

1. Perform eDNA sampling of seawater at 4 sampling points along each of 7 trawl tracks at 2018 Biosurvey sampling stations in the Ports of Los Angeles and Long Beach, immediately prior to benthic trawls at those sites.
2. Analyze the eDNA samples to get an inventory of unique sequences from the samples, targeting fish and benthic invertebrates.
3. Perform bioinformatic analysis of the sequence results to get the best possible estimates of the taxonomic composition and diversity at each of the sites.
4. Compare the eDNA diversity results with the conventional trawl and overall 2018 Biosurvey program diversity results.
5. Use benthic marine invertebrate voucher specimens from the Port biodiversity survey and other sources to generate at least 200 reference barcode sequences for local marine invertebrates.

## 2 MATERIALS AND METHODS

### 2.1 Overview of environmental DNA sampling and processing

Figure 2-1 shows an overview of the steps followed to sample and process the environmental DNA samples in this study. The detailed methods needed to reproduce the results of this study are available in Appendix II.

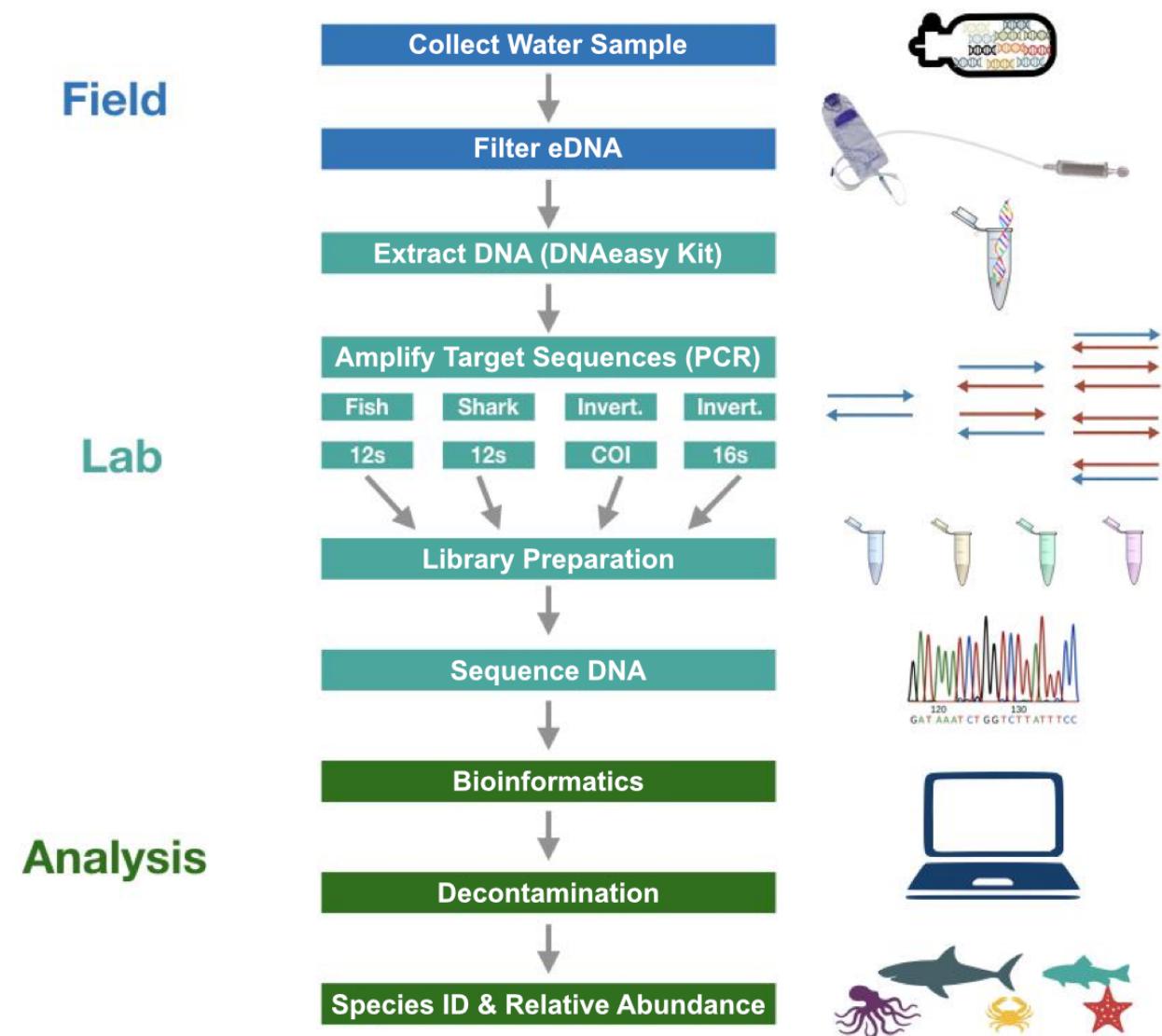


Figure 2-1. Flow chart of eDNA sample processing.

### 2.2 Seawater sampling of environmental DNA

To assess Ports of Los Angeles and Long Beach marine fish and invertebrate communities using eDNA, we sampled eDNA at seven 2018 Biosurvey stations (Appendix III, panels a–g) on 20–21 August 2018 in San Pedro Bay from the R/V Yellowfin, operated by the Southern California Marine

Institute (Table 2-1). Samples were taken at standardized sampling stations used for biological diversity sampling by the Port of Los Angeles and the Port of Long Beach. At each station, water was sampled at 4 locations along a 150 m transect. At each location along a station transect, 3 replicate 1 L water samples were collected ( $3 \times 4 = 12$  total samples per station transect plus a blank control) using three 5 L Niskin bottles (General Oceanics), approximately 2 m above the seafloor (depth ranged from 6–23 m). At one location, an extra set of three samples was taken at the surface (approximately 1 m depth) in addition to the regular set of samples above the seafloor.

Sampling Level	Sample Enumeration	Total Number
Stations	LA3, LA6, LA11, LA14, LB1, LB2, LB14	7 stations
Locations (4 per station)	(4 $\times$ 7 stations) + 1 surface location	29 locations
Samples (3 per location)	(3 $\times$ 29 locations) + (1 blank $\times$ 7 stations)	87 samples + 7 blanks

*Table 2-1. Sampling design for eDNA seawater sampling.*

*At one location at station LB1, an extra set of three samples was taken at the surface in addition to the regular set of three samples from 2 m above the seafloor. Station locations are shown in Appendix III, panels a–g.*

Upon arrival at each location, sampling bottles were flushed continuously with surface water from the location for a minimum of five minutes before deployment. Samples were taken serially, separately deploying each Niskin bottle while the boat was stationary. Immediately upon retrieval of each Niskin sample bottle, it was removed from the hydrographic line and placed in a rack on deck and immediately tapped into a 1 L Kangaroo Enteral Feeding Gravity Bag (Covidien PLC, Dublin, Ireland), the first 0.5 L of each sample was used to rinse the bag and then discarded, the bag was then filled with 1 L of sample water and gravity filtered through 0.22 $\mu$ m Sterivex filters (Millipore Corp. Burlington, MA, USA) (Miya et al. 2015, Port et al. 2016, Curd et al. 2019). Filtration was carried out in the ship cabin, shielded from sunlight. After filtration, excess water was removed from the filter cartridge with a sterile 3 mL syringe and both ends were secured with Luer Lock caps (B. Braun Melsungen AG, Melsungen, Germany). Each dried Sterivex filter was placed into an individual sterile Whirl-Pak bag (Nasco, Fort Atkinson, WI, USA) and then placed in a cooler with dry ice and ice packs (Curd et al. 2019). As a negative control, at each station we processed a field blank that consisted of 1 L of UltraPure DNase/RNase Free Distilled Water (Invitrogen, Carlsbad, CA) following standard methods (Goldberg et al. 2016, Kelly et al. 2018). All samples were transferred into a -20° F freezer (Miya et al. 2015) within 6 hours of sample collection and extracted within 1 week of the field sampling.

## 2.3 Trawl and other 2018 Biosurvey sampling

Multiple surveys were conducted during the 2018 Biosurvey, which occurred between April 2018 and March 2019. The eDNA sampling for this study was conducted in August 2018 in conjunction with the daytime benthic otter trawls (Appendix III, panels a–g). At each of the sites sampled for eDNA, as soon as seawater sampling along the trawl track was complete (described above), the location was trawled with a 7.6 m semi-balloon otter trawl aboard the R/V Early Bird II. In all cases, the trawl followed the water sampling within 15–90 minutes.

In order to compare the broader coverage provided by eDNA sampling with biodiversity from the range of 2018 Biosurvey gear types, data from other surveys (e.g., lampara for pelagic fishes, diver surveys for hard-substrate associated algae and invertebrates, and benthic infauna sampling from sediment grabs) were included (Appendix III, panels a–g). In this document these data will be referred to as the 2018 Biosurveys. In order to account for seasonal variation, data from the same season as the eDNA sampling were used to compare how well the detection of fish, algae, and invertebrate species correlated between the two sampling methods.

Detailed methodology on the Biosurveys program, including details of the otter trawl sampling that paralleled the eDNA sampling, is available in the 2018 Biosurvey report (Wood Environment and Infrastructure Solutions, Inc. 2020).

## 2.4 Contamination precautions

To minimize the possibility of contamination (which can cause false positives), physical contamination controls were used in the field and laboratory, following procedures of Goldberg et al. (2016). Personnel were gloved, and supplies and facilities were sterilized to remove or destroy residual DNA. Full methods are in Appendix I.

## 2.5 Environmental DNA extractions, amplification, and sequencing

At University of California Los Angeles (UCLA), all eDNA filter extractions and PCR preparations were conducted in an AirClean 600 PCR Workstation (Creedmoor, NC, USA) in a clean room dedicated to DNA extractions. There were negative controls (extraction blanks and PCR blanks) for all steps in the sample processing. Full methods are in Appendix I.

Sterivex filters were extracted with a DNAeasy Tissue and Blood Qiagen Kit (Spens et al. 2017). DNA was amplified with primer sets using Illumina Nextera adapter modifications for Nextera XT (indexes Product Number FC-131-2001 through -2004; 5200 Illumina Way, San Diego, CA 92122). The four primer sets used were: (1) MiFish Universal Teleost 12S primer (176bp) (Miya et al. 2015); (2) MiFish Universal Elasmobranch 12S primer (186bp) (Miya et al. 2015); (3) metazoan COI primers (313bp) (Leray et al. 2013); and (4) metazoan 16S primers (~115bp) (Kelly et al. 2016). All PCRs included a negative control, where molecular grade water replaced the DNA template. All PCR products were run on 2% agarose gels to assess amplification success and correct product size.

To prepare PCR products for sequencing, triplicate PCR reactions were pooled using 5 µL volume from each PCR. Pooled samples were cleaned using Serapure magnetic beads (Faircloth and Glenn 2014) to isolate target DNA from unused primers and DNA polymerase enzymes. Cleaned PCR product concentrations were quantified using the high sensitivity Quant-iT™ dsDNA Assay Kit (Thermofisher Scientific, Waltham, MA, USA) on a Victor3™ plate reader (Perkin Elmer Waltham, MA, USA). All indexed PCR products were run on 2% agarose gels to assess successful PCR and correct DNA fragment size. Resulting libraries were bead cleaned and quantified as described above, then sequenced at the Technology Center for Genomics & Bioinformatics (UCLA, CA, USA).

## 2.6 Bioinformatic analysis and taxonomic determination

The *Anacapa Toolkit* (Curd et al. 2019) was used for amplicon sequence variant (ASV) parsing, taxonomic assignment, and quality control. Taxonomic assignments of eDNA metabarcoding reads require the sequence reference library to be as comprehensive as possible. To generate libraries for fish and invertebrates, we queried NCBI GenBank on 28 May 2019 for sequences that overlapped our primer regions. Four libraries were constructed: 1) 12S for bony fish; 2) 12S for rays, and sharks; 3) COI; and 4) 16S for metazoans including invertebrates. Sequences were processed using the default parameters and then assigned taxonomy using reference libraries generated using CRUX (Constructing Reference libraries Using eXisting tools) (Curd et al. 2019). The quality control step of the *Anacapa Toolkit* trims extraneous adapter sequences used to identify each unique sample, removes low quality reads, and sorts reads by metabarcode primer sequence.

The ASV parsing step in the *Anacapa Toolkit* uses DADA2 (Callahan et al. 2016) to derePLICATE the metabarcodes. ASVs are a novel solution to identifying biologically informative unique sequences in metabarcoding samples that replaces the operational taxonomic unit (OTU) framework. Unlike OTUs, which cluster sequences using an arbitrary sequence similarity (e.g. 97%, which corresponds to a gross average genetic sequence difference across vertebrates), ASVs are unique sequence reads determined using Bayesian probabilities of known sequencing error from Illumina platforms. These unique sequences can differ by as little as two base pairs, providing improved taxonomic resolution and an increase in observed diversity (Callahan et al. 2016, Amir et al. 2017).

Taxonomy is assigned based on the lowest common ancestor of multiple reference library matches for each query sequence. The reliability of each taxonomic assignment is then evaluated through bootstrap confidence scores (Gao et al. 2017). Scores are based on Bayesian posterior probability which quantify the similarity of reference barcode sequences to the query sequence. The higher the similarity between the database barcode sequence and the query sequence, the greater the contribution to the taxonomic assignment of the query. Ultimately, this method provides a strong probabilistic basis for evaluating taxonomic assignments with bootstrap confidence scores.

## 2.7 Bioinformatic decontamination

After processing the raw sequence reads through the *Anacapa Toolkit*, the resulting species community tables were transferred into R (R Core Team 2016) for subsequent decontamination and downstream data analysis. The raw species community table needs to be decontaminated to eliminate potential field contamination, lab contamination, and sequence index hopping (Goldberg et al. 2016, Costello et al. 2018). Field and lab contamination can arise from traces of non-target DNA introduced by inadequate sterile procedures, careless laboratory work, and commercial reagents, particularly enzymes generated from living organisms (Goldberg et al. 2016). Sequence index hopping occurs when the DNA index tag used to label each unique sample chemically swaps with the DNA index tag of another sample, leading to cross-contamination of species between samples. Rigorous sterile procedures were followed and data was sequenced on platforms that limit index hopping.

Decontamination was a three step procedure (Kelly et al. 2018 [supplemental methods], 2019 [supplemental methods], Gallego 2019): 1) removal of sequences arising from index hopping, 2) removal of sequences arising from negative controls, 3) removal of low confidence detections through site occupancy modeling. Briefly, the first step was accomplished by estimating the rate of index hopping of each ASV into negative controls and used as a baseline for removing likely molecular recombination of index sequences. Second, the likely origin of each ASV was identified from either samples or control by comparing the total number of reads, prevalence (number of samples in which reads occurred), the proportion of reads, and proportion of prevalence of each ASV in both samples and controls. These metrics were compared for a given ASV in samples and controls and any ASVs with metrics greater in controls than samples were eliminated. Third, we estimated the probability that a species was a true detection based on the prevalence of detection across multiple replicate samples at a given site (Chambert et al. 2018). Together, this highly conservative decontamination approach ensures high confidence in the results.

Following taxonomic assignments, fish species identified from eDNA approaches were checked against known distributions in FishBase (Froese and Pauly 2019) to provide ground truthing and ecological context. This information was then used to determine whether a given species detected by eDNA methods was likely to be present in the environment or is an artifact of the bioinformatic assignment process. Species that were considered unlikely to be present included those that were freshwater, tropical, common in the aquarium trade, seafood species whose ranges do not include central/southern California or Mexico, and species whose known distribution does not include the Pacific Ocean side of North America.

Because of the much greater diversity of invertebrates, the accompanying greater prevalence of taxonomic naming discrepancies, and the notably poorer proportion of expected species that were available in sequence reference databases, invertebrate taxonomic estimates were treated in quantitative fashion, and were not individually assessed.

## 2.8 Normalizing environmental DNA reads for abundance estimation

Data were transformed into an “eDNA index” (Kelly et al. 2019). This metric assumes that PCR biases originate from template-primer interactions which remain constant across eDNA samples. Intuitively, therefore, it should be possible to infer relative abundance within taxa across samples by using appropriate normalizations of each sample’s read counts. The eDNA index transformation is conducted by first normalizing all reads for a particular sequence by the total number of reads in each sample, then scaling those proportions to the largest observed proportion for that sequence across all samples. This results in a sequence-specific (species-specific) scaling between 0 to 1, where 1 is the sample with the highest number of reads for a given species and 0 is the least.

## 2.9 Comparisons of taxon occurrences

In order to compare species identified with eDNA metabarcoding and conventional survey methods, taxonomic rank (phylum, class, order, family, genus, and species) of organisms identified by Biosurvey methods and eDNA metabarcoding were compared in R (R Core Team 2016). The 2018

Biosurvey species list was compared to the eDNA results. Biosurvey and eDNA results were also compared to the CAL-NEMO invasive species database (Fofonoff et al. 2019).

Data analysis used the *phyloseq* (McMurdie and Holmes 2013) and *ranacapa* (Kandlikar et al. 2018) R packages to import and manipulate the eDNA metabarcoding data in R. In order to investigate the relationship between eDNA index scores and trawl biomass and relative abundance, linear regressions between both metrics were fit in R, and  $R^2$  and equations for each fit were reported.

Comparisons of the total number of species found at each site from eDNA metabarcoding analysis were run using Analysis of Variance (ANOVA) and Tukey post-hoc tests using the R package *vegan* (Oksanen et al. 2019).

Pairwise similarity of all 87 eDNA samples was calculated using the Bray-Curtis similarity index through the *vegan* package. The analysis apportioned variation between the pairwise similarity of each sample against site, station, data collected, and volume filtered using a PERMANOVA and Betadisp analysis. Apportioned variance was visualized using the *treemapify* package (Wilkins 2019) and a non-metric multidimensional scaling (NMDS) ordination through *phyloseq* and *vegan* packages. An NMDS plot collapses information about comparative ranking of species occurrences into two dimensions, giving insight into similarities between samples, locations, and sites.

It is intuitively clear that sampling a community more times will yield higher estimates of total species richness. Common species are collected early and easily; more sampling yields higher species counts by increasing the capture of rarer species. Eventually, though, sampling reaches “saturation”, where increased sampling no longer captures more species. Calculation of “rarefaction curves” formalizes this relationship between sample number and estimated species richness, giving insight into what density of sampling is needed to capture a high fraction of the community’s total species richness (Heck et al. 1975). Species rarefaction curves were calculated to analyze the number of additional species found for each additional replicate across each site using the *iNext* (Chao et al. 2014, Hsieh et al. 2019) and *vegan* R packages.

## 2.10 Generation of barcode reference sequences

Invertebrate voucher specimen sampling (Department of Fish and Wildlife Permit CA-004668) was not limited to benthic trawls, but also included other sampling methods (e.g. beach seines and bottom grabs) from Los Angeles County localities. Specimens were identified to the level of species. Genomic DNA was isolated using either: 1) the DNeasy® Blood and Tissue Kit (Qiagen, Valencia, CA, USA) following the manufacturer’s protocol or 2) using an AutoGenprep 965 using the AutoGen standard Mouse Tail (Animal Tissue) protocol (Holliston, MA, USA). For fishes, MiFish 12S primers were used (Miya et al. 2015), and for invertebrates, metazoan COI (Geller et al. 2013) primers were used. PCR products were purified using ExoSAP-IT (Affymetrix, Cleveland, OH, USA) and sequenced in both primer directions. Forward and reverse sequences were trimmed at the ends and assembled into contigs in Geneious version 8.1.9 (<http://www.geneious.com>) (Kearse et al. 2012) and aligned using MUSCLE (Edgar 2004). All reference sequences will be shared publicly at the Barcode of Life Database (BOLD, <http://boldsystems.org>) and (via BOLD) to GenBank.

### 3 RESULTS

#### 3.1 Seawater sampling of environmental DNA

Table 3-1 enumerates the seawater samples that were taken for eDNA analysis. The volume filtered reported for each sample reflects the maximum volume (up to 1 L) that could practically be filtered through the Sterivex filter in the time available on board the sampling vessel.

Station	Date Sampled	Location	Sample	Volume Filtered (mL)	Station	Date Sampled	Location	Sample	Volume Filtered (mL)
LA3	8/20/2018	1	1	900	LB2	8/21/2018	1	1	820
			2	990				2	850
			3	800				3	800
		2	1	950			2	1	680
			2	880				2	800
			3	980				3	700
		3	1	910			3	1	800
			2	850				2	800
			3	1000				3	950
		4	1	850			4	1	750
			2	890				2	700
			3	890				3	950
		Blank	Blank	900			Blank	Blank	1000
LA11	8/20/2018	1	1	970	LB1	8/21/2018	Surface	1	920
			2	990				2	900
			3	910				3	910
		2	1	1000			1	1	700
			2	940				2	800
			3	1000				3	690
		3	1	995			2	1	720
			2	930				2	780
			3	1000				3	730
		4	1	930			3	1	720
			2	995				2	700
			3	960				3	990
		Blank	Blank	1000			4	1	780
LA14	8/20/2018	1	1	890				2	750
			2	900				3	800
			3	910				Blank	Blank
		2	1	910			1	1	995
			2	1000				2	1000
			3	1000				3	995
		3	1	820				2	980
			2	810				2	990
			3	900				3	1000
		4	1	860			3	1	995
			2	920				2	960
			3	900				3	1000
		Blank	Blank	1000			4	1	1000
LA6	8/20/2018	1	1	1000				2	1000
			2	1000				3	1000
			3	1000				Blank	Blank
		2	1	985			1	1	1000
			2	1000				2	1000
			3	995				3	1000
		3	1	1000				Blank	Blank
			2	1000				Blank	Blank
			3	1000				Blank	Blank
		4	1	1000				Blank	Blank
			2	1000				Blank	Blank
			3	1000				Blank	Blank
		Blank	Blank	1000				Blank	Blank

Table 3-1. Seawater samples taken at each location at each station for eDNA analysis.

The “Surface” entries at station LB1 reflect a one-time collection of three samples at the surface (approximate 1 m depth) in addition to the standard sampling at all locations.

### 3.2 Sequencing

For eDNA, the Anacapa Toolkit identified 12,722 ASVs from 21,461,855 reads (individual sequences returned from the MiSeq sequencing machine) from 87 samples across 4 barcodes (Table 3-2). Despite rigorous sterile field and lab procedures, a low level of contamination occurred, including the detection of California fish species in negative controls. Stringent bioinformatic decontamination procedures removed all contaminated sequences. After decontamination procedures, we retained 12,384 ASVs out of 11,379,704 reads for subsequent data analysis.

Barcode	Samples with Successful Data	Raw Reads	Raw ASVs	Post Decontamination Reads	Post Decontamination ASVs
16S	82	7,385,650	2,309	5,251,929	2,187
CO1	87	3,520,548	7,058	2,054,308	6,944
12S ELAS	87	3,306,204	1,299	1,528,931	1,257
12S MIU	87	7,249,453	2,056	2,544,536	1,996
<b>Total</b>		<b>21,461,855</b>	<b>12,722</b>	<b>11,379,704</b>	<b>12,384</b>

Table 3-2. Summary of sequences generated from eDNA metabarcoding.

### 3.3 Reference barcode generation

As described above, taxa for which there is no reference sequence cannot be correctly identified using eDNA. Therefore, to augment publicly-available bony and cartilaginous fish reference sequences, additional fish barcode references were generated for this project, based on fish species lists generated by the 2018 Biosurvey (Table 3-3).

Reference Sequences Generated	Scientific Name	Common Name
2	<i>Paralabrax clathratus</i>	Kelp Bass
1	<i>Paralabrax maculatofasciatus</i>	Spotted Sand Bass
5	<i>Paralabrax nebulifer</i>	Barred Sand Bass
6	<i>Paralichthys californicus</i>	California Halibut
1	<i>Phanerodon furcatus</i>	White Surfperch
2	<i>Pleuronichthys ritteri</i>	Spotted Turbot
5	<i>Seriphis politus</i>	Queenfish
1	<i>Sympodus atricaudus</i>	California Tonguefish
1	<i>Urolophus halleri</i>	Round Stingray

Table 3-3. Fish barcode references created for this project.

For invertebrates, public databases are far less complete than they are for vertebrates. Hence for this project, a goal of sequencing an additional 200 reference barcode sequences for local invertebrates was proposed and met, with a total of 298 new barcode references created (Appendix I).

Vouchered and databased specimens (from which the barcode references were generated) are permanently housed in the NHMLA collections.

### **3.4 Survey comparisons for fish**

There is a range of comparisons that can be made of the taxa reported by the various survey methods. The most direct (and most restrictive) comparison between methods is to compare taxa observed in the eDNA survey with taxa observed in the parallel trawls on the same day. Because of the integrative nature of eDNA in water samples, it is also appropriate to compare taxa observed in the eDNA survey with all taxa observed in the full 2018 Biosurvey program. Because reference barcode coverage for fish is dramatically better than for invertebrates, fish and invertebrate comparisons are treated separately.

All fish species observed during the trawls made in parallel with the eDNA sampling were observed by eDNA (Table 3-4). The observational overlap was achieved only after this project added an additional 16 fish reference barcodes to fill in taxa with missing reference barcodes; without those additional barcode references, those taxa could not have been detected by eDNA.

Station		All Stations 6–25		LA3 6		LA6 17.5		LA11 25		LA14 7.5		LB1 13		LB2 7		LB14 18	
Scientific Name	Common Name	Trawl	eDNA	Trawl	eDNA	Trawl	eDNA	Trawl	eDNA	Trawl	eDNA	Trawl	eDNA	Trawl	eDNA	Trawl	eDNA
<i>Citharichthys xanthostigma</i>	Longfin Sanddab	1	✓		✓			1	✓						✓		
<i>Cymatogaster aggregata</i>	Shiner Surfperch	6	✓		✓		✓	6	✓		✓				✓		✓
<i>Engraulis mordax</i>	Northern Anchovy	200	✓		✓	2	✓		✓		✓		✓	198	✓		✓
<i>Genyonemus lineatus</i>	White Croaker	115	✓		✓		✓	47	✓		✓	30	✓	38	✓		✓
<i>Hypsopsetta guttulata</i>	Diamond Turbot	1	✓		✓		✓		✓		✓		✓		✓	1	✓
<i>Paralabrax maculatofasciatus</i>	Spotted Sand Bass	1	✓		✓		✓		✓	1	✓		✓		✓		✓
<i>Paralabrax nebulifer</i>	Barred Sand Bass	14	✓	4	✓	4	✓		✓	1	✓		✓		✓	5	✓
<i>Paralichthys californicus</i>	California Halibut	8	✓	3	✓	1	✓		✓		✓		✓	1	✓	3	✓
<i>Phanerodon furcatus</i>	White Surfperch	1	✓	1	✓		✓		✓		✓			✓		✓	
<i>Pleuronichthys ritteri</i>	Spotted Turbot	1	✓		✓		✓		✓		✓		✓	1	✓		✓
<i>Porichthys myriaster</i>	Specklefin Midshipman	1	✓		✓		✓		✓			1	✓		✓		✓
<i>Scorpaena guttata*</i>	Spotted Scorpionfish	1	✓		✓			1	✓			✓					✓
<i>Seriphus politus</i>	Queenfish	36	✓		✓	36	✓		✓		✓		✓		✓		✓
<i>Syphurus atricaudus</i>	California Tonguefish	2	✓		✓	1	✓		✓		✓		✓		✓	1	✓
<i>Synodus lucioceps</i>	California Lizardfish	60	✓	16		1	✓	14	✓			26	✓		✓	3	✓
<i>Urolophus halleri</i>	Round Stingray	1	✓	1	✓		✓		✓		✓		✓		✓		✓
<i>Xystreurus liolepis</i>	Fantail Sole	3	✓		✓		✓	3	✓		✓		✓		✓		✓
<b>Total Taxa</b>		<b>17</b>	<b>17</b>	<b>5</b>	<b>16</b>	<b>6</b>	<b>15</b>	<b>6</b>	<b>17</b>	<b>2</b>	<b>13</b>	<b>3</b>	<b>14</b>	<b>4</b>	<b>16</b>	<b>5</b>	<b>16</b>

\* Note that the eDNA survey matched to *Scorpaena pepo*, congeneric to *S. guttata*. This likely represents the closest available match for *S. guttata* sequences and is interpreted as representing *S. guttata* for this study.

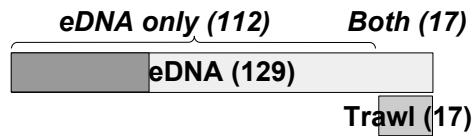
Table 3-4. Enumeration of fish taxa observed in parallel trawl and eDNA samples.

Numbers in the “Trawl” column are counts; a checkmark in the eDNA column denotes detection of the taxon by eDNA.

It is also possible to compare the trawl results with all fish taxa detected by eDNA. Figure 3-1 compares the number of fish taxa observed in the trawls made parallel with the eDNA survey with all fish species detected by eDNA. A total of 129 fish taxa were detected by eDNA. Of the 129, 44 taxa resolve to DNA matches that represent fish taxa but seem unlikely to represent correct assignments to a local species (denoted with the darker color in Figure 3-1), leaving 85 fish taxa detected only by eDNA that seem reasonable based on species and geography.

*Figure 3-1. Counts of fish taxa observed in the eDNA data and trawl sampling.*

*The area where the bars overlap shows species detected by both methods. The darker end of the eDNA species counts indicates unique fish taxa that appear unlikely to be correctly taxonomically matched to local species.*



A broader comparison is between the eDNA observations and all fish taxa observed across all Biosurvey dates and methodologies. Figure 3-2 is a summary diagram that enumerates the counts of taxa from eDNA, Biosurvey, and how they compare.

*Figure 3-2. Counts of fish taxa observed in the eDNA data and Biosurvey.*

*The area where the bars overlap shows species detected by both methods. The darker end of the eDNA species counts indicates unique fish taxa that appear unlikely to be correctly taxonomically matched to local species.*

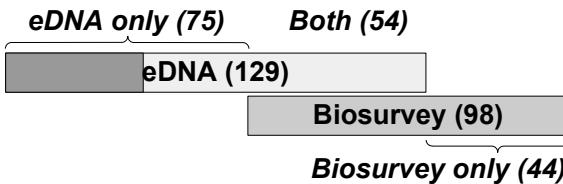


Table 3-5 lists the 54 fish taxa observed in both the eDNA and full Biosurvey observations. Table 3-6 lists the 44 fish taxa that were observed in the full Biosurvey but did not have a match in the eDNA observations.

*Table 3-5. Fish taxa observed in both eDNA data and Biosurvey.*

Scientific Name	Common Name
<i>Acanthogobius flavimanus</i>	Yellowfin Goby
<i>Amphistichus argenteus</i>	Barred Surfperch
<i>Atherinops affinis</i>	Topsmelt
<i>Atherinopsis californiensis</i>	Jacksmelt
<i>Brachyistius frenatus</i>	Kelp Perch
<i>Cheilotrema saturnum</i>	Black Croaker
<i>Chromis punctipinnis</i>	Blacksmith
<i>Citharichthys sordidus</i>	Pacific Sanddab
<i>Citharichthys stigmatus</i>	Speckled Sanddab
<i>Citharichthys xanthostigma</i>	Longfin Sanddab
<i>Clevelandia ios</i>	Arrow Goby
<i>Clinocottus analis</i>	Woolly Sculpin
<i>Cymatogaster aggregata</i>	Shiner Surfperch
<i>Rhacochilus vacca</i>	Pile Surfperch
<i>Embiotoca jacksoni</i>	Black Surfperch
<i>Engraulis mordax</i>	Northern Anchovy
<i>Genyonemus lineatus</i>	White Croaker
<i>Gillichthys mirabilis</i>	Longjaw Mudskipper
<i>Girella nigricans</i>	Opaleye
<i>Halichoeres semicinctus</i>	Rock Wrasse
<i>Hermosilla azurea</i>	Zebra-perch sea chub
<i>Heterodontus francisci</i>	Horn Shark
<i>Hypsoblennius gentilis</i>	Bay Blenny
<i>Hypsopsetta guttulata</i>	Diamond Turbot
<i>Hypsypops rubicundus</i>	Garibaldi
<i>Lepidogobius lepidus</i>	Bay Goby
<i>Leuresthes tenuis</i>	California Grunion
<i>Medialuna californiensis</i>	Halfmoon
<i>Myliobatis californica</i>	Bat Ray
<i>Ophidion scriptpsi</i>	Basketweave Cusk-eel
<i>Oxyjulis californica</i>	Senorita
<i>Paralabrax clathratus</i>	Kelp Bass
<i>Paralabrax maculatofasciatus</i>	Spotted Sand Bass
<i>Paralabrax nebulifer</i>	Barred Sand Bass
<i>Paralichthys californicus</i>	California Halibut
<i>Phanerodon furcatus</i>	White Surfperch
<i>Pleuronichthys ritteri</i>	Spotted Turbot
<i>Porichthys myriaster</i>	Specklefin Midshipman
<i>Porichthys notatus</i>	Plainfin Midshipman
<i>Raja inornata</i>	California Skate
<i>Scomber japonicus</i>	Pacific Mackerel
<i>Sebastes paucispinis</i>	Bocaccio
<i>Semicossyphus pulcher</i>	California Sheephead
<i>Seriola lalandi</i>	Yellowtail amberjack
<i>Seriphis politus</i>	Queenfish
<i>Stereolepis gigas</i>	Black Sea Bass
<i>Syphurus atricaudus</i>	California Tonguefish
<i>Syngnathus leptorhynchus</i>	Bay Pipefish
<i>Synodus lucioceps</i>	California Lizardfish
<i>Trachurus symmetricus</i>	Jack Mackerel
<i>Tridentiger trigonocephalus</i>	Chameleon Goby
<i>Urolophus halleri</i>	Round Stingray
<i>Xystreurus liolepis</i>	Fantail Sole
<i>Zaniolepis latipinnis</i>	Longspine Combfish
<b>Total</b>	<b>54</b>

*Table 3-6. Fish taxa observed only in Biosurvey data.*

Species Name	Common Name
<i>Anchoa compressa</i>	Deepbody Anchovy
<i>Anchoa delicatissima</i>	Slough Anchovy
<i>Anisotremus davidsonii</i>	Sargo
<i>Caulolatilus princeps</i>	Ocean Whitefish
<i>Cosmocampus arctus</i>	Snubnose Pipefish
<i>Cryptotrema corallinum</i>	Deep-Water Blenny
<i>Fodiator acutus</i>	Sharpchin Flyingfish
<i>Gobiesox rhessodon</i>	California Clingfish
<i>Gymnothorax mordax</i>	Moray Eel
<i>Heterostichus rostratus</i>	Giant Kelpfish
<i>Hippocampus ingens</i>	Pacific Seahorse
<i>Hypsoblennius jenkinsi</i>	Mussel Blenny
<i>Icelinus quadriseriatus</i>	Yellowchin Sculpin
<i>Leptocottus armatus</i>	Staghorn Sculpin
<i>Lyopsetta exilis</i>	Slender Sole
<i>Lythrypnus dalli</i>	Blue-Banded Goby
<i>Lythrypnus zebra</i>	Zebra Goby
<i>Menticirrhus undulatus</i>	California Corbina
<i>Micrometrus minimus</i>	Dwarf Surfperch
<i>Neoclinus uninotatus</i>	Onespot Fringehead
<i>Odontopyxis trispinosa</i>	Pygmy Poacher
<i>Paraclinus integrifinnis</i>	Reef Finspot
<i>Parophrys vetulus</i>	English Sole
<i>Peprilus simillimus</i>	Pacific Butterfish
<i>Platyrrhinoidis triseriata</i>	Thornback Ray
<i>Pleuronichthys guttulata</i>	Diamond Turbot
<i>Pleuronichthys verticalis</i>	Hornyhead Turbot
<i>Rhinobatos productus</i>	Shovel-nose guitarfish
<i>Rhinogobiops nicholsii</i>	Blackeye Goby
<i>Rusarius creaseri</i>	Roughcheek Sculpin
<i>Sardinops sagax caeruleus</i>	Pacific Sardine
<i>Scorpaena guttata</i>	Spotted Scorpionfish
<i>Sebastes auriculatus</i>	Brown Rockfish
<i>Sebastes carnatus</i>	Gopher Rockfish
<i>Sebastes miniatus</i>	Vermillion Rockfish
<i>Syngnathus arctus</i>	Snubnose Pipefish
<i>Syngnathus aulicus</i>	Barred Pipefish
<i>Sphyraena argentea</i>	Barracuda
<i>Torpedo californica</i>	Pacific Electric Ray
<i>Trachinotus paitensis</i>	Paloma Pompano
<i>Typhlogobius californiensis</i>	Blind Goby
<i>Umbrina roncador</i>	Yellowfin Croaker
<i>Xenistius californiensis</i>	Salema
<i>Zaniolepis frenata</i>	Shortspine Combfish
<b>Total</b>	<b>44</b>

Table 3-7 and Table 3-8 show the remaining 75 eDNA observations of fish that were not found in the Biosurvey. Based on an initial evaluation, these have been separated into 31 taxa that appear to be reasonable as local species (Table 3-7), and 44 taxa whose closest sequence match appears unlikely to be a direct match to a local taxon (Table 3-8).

*Table 3-7. Fish taxa observed only in eDNA data deemed reasonable as correct matches.*

Scientific Name	Common Name	Comments
<i>Alopias vulpinus</i>	Thresher shark	
<i>Carcharodon carcharias</i>	Great white shark	
<i>Chilara taylori</i>	Spotted cusk-eel	
<i>Chitonotus pugetensis</i>	Roughback sculpin	
<i>Cololabis saira</i>	Pacific saury	
<i>Coryphaena hippurus</i>	Common dolphinfish	
<i>Decapterus macrosoma</i>	Shortfin scad	
<i>Etrumeus teres</i>	Red-eye round herring	
<i>Girella simplicidens</i>	Gulf opaleye	Mahi-mahi
<i>Gobiesox maeandricus</i>	Northern clingfish	Tropical, seen in Mexico but not reported in US. Range expansion?
<i>Gymnura marmorata</i>	California butterfly ray	Subtropical, NW Atlantic. One report from Monterey Bay, CA 2005
<i>Icelinus burchami</i>	Dusky sculpin	Gulf of CA usually. Range expansion? Possible misannotation of our local opaleye species
<i>Icelinus filamentosus</i>	Threadfin sculpin	
<i>Isurus oxyrinchus</i>	Shortfin mako shark	
<i>Katsuwonus pelamis</i>	Skipjack tuna	
<i>Melichthys niger</i>	Black triggerfish	Possible range expansion from San Diego, or mismatch to finescale triggerfish
<i>Mugil cephalus</i>	Flathead grey mullet	
<i>Myrophis vafer</i>	Pacific worm eel	
<i>Oncorhynchus keta</i>	Chum salmon	Possible seafood influence, natural range as far south as LA/SD seems unlikely.
<i>Peprilus medius</i>	Pacific harvestfish	
<i>Pleuronichthys decurrens</i>	Curlfin sole	
<i>Prionace glauca</i>	Blue Shark	
<i>Pteroplatytrygon violacea</i>	Pelagic stingray	
<i>Sarda chiliensis</i>	Eastern Pacific bonito	Subtropical, noted in Baja. Range expansion?
<i>Sardinops melanostictus</i>	South American pilchard	Lineage (caeruleus) known to CA
<i>Scopeloberyx robustus</i>	Longjaw bigscale	
<i>Scorpaenichthys marmoratus</i>	Cabezon	
<i>Semicossyphus pulcher</i>	Sheephead	
<i>Thaumatichthys pagidostomus</i>	Anglerfish	
<i>Thunnus alalunga</i>	Albacore tuna	
<i>Triakis semifasciata</i>	Leopard shark	
<b>Total</b>	<b>31</b>	

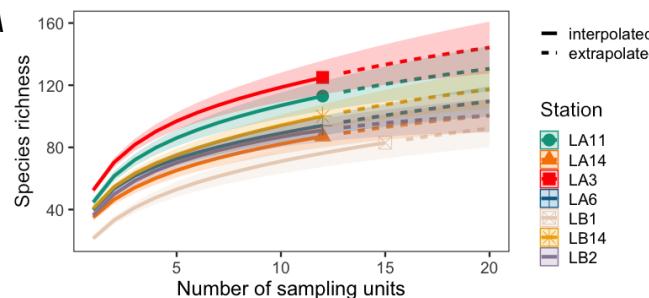
**Table 3-8. Additional fish taxa observed only in eDNA data.**

These represent unique taxa (based on sequence), but may not be taxonomically correct matches to local taxa.

Scientific Name	Common Name	Comments
<i>Acanthopagrus sp. YI-2011a</i>	Seabream sp.	Undetermined. Seabream spp not reported in North America, seafood species. Unlikely
<i>Agamyxis albomaculatus</i>	Spiny cat-fish	Freshwater, tropical. Aquarium? Unlikely
<i>Albula glossodonta</i>	Roundjaw bonefish	Tropical, not seen in US although <i>Albula vulpes</i> is in CA. Maybe misannotation. Unlikely
<i>Apogon imberbis</i>	Cardinal fish	Atlantic subtropical, Mediterranean. Aquarium? Unlikely
<i>Astropteryx semipunctata</i>	Starry goby	Indopacific tropical. Unlikely
<i>Cobitis choii</i>	Choi's spiny loach	freshwater, subtropical. Aquarium? Unlikely
<i>Engraulis encrasicolus</i>	European anchovy	Europe and Africa, not reported in US. Maybe misannotation Unlikely
<i>Equetus lanceolatus</i>	Jack-knifefish	Tropical, Caribbean. Unlikely
<i>Gymnogobius heptacanthus</i>	Goby	Asia, not reported in US. Unlikely
<i>Gymnogobius opperius</i>	Goby	Freshwater, Asia/Russia. Unlikely
<i>Gymnogobius urotaenia</i>	Goby	Freshwater, Asia/Russia. Unlikely
<i>Gymnothorax meleagris</i>	Turkey moray	Tropical, Indopacific. Maybe misannotation. Unlikely
<i>Luciogobius guttatus</i>	Flat-headed goby	Subtropical, Asia. Not reported in North America. Unlikely
<i>Lutjanus campechanus</i>	Northern red snapper	Subtropical, Caribbean (FL). Seafood. Unlikely
<i>Malacanthus brevirostris</i>	Quakerfish	Tropical, as close as Hawaii and central America. Unlikely
<i>Melanogrammus aeglefinus</i>	Haddock	N Atlantic, very common in seafood and animal feed. Unlikely
<i>Myliobatis aquila</i>	Common eagle ray	Europe and Africa, not reported in US. Maybe misannotation Unlikely
<i>Neoclinus nudus</i>	Blenny	Asia. Not reported in North America. Unlikely
<i>Odontesthes bonariensis</i>	Argentinian silverside	FW/brackish, South America. Introduced in Asia, Europe. No North America reports. Unlikely
<i>Odontesthes incisa</i>	Silverside	SE South America, no reports in North America. Unlikely
<i>Odontesthes smitti</i>	Silverside	SE South America, no reports in North America. Unlikely
<i>Odontesthes sp. Odsp-001</i>	Silverside	Undetermined, but likely just a misannotation of the above. Unlikely
<i>Odontesthes sp. TJX-2014</i>	Silverside	Undetermined, but likely just a misannotation of the above. Unlikely
<i>Oncorhynchus tshawytscha</i>	Chinook salmon	Range extends to N California, possible seafood influence. Unlikely
<i>Opistognathus evermanni</i>	Jawfish	Asia. Not reported in North America. Unlikely
<i>Opistognathus iyonis</i>	Jawfish	Asia. Not reported in North America. Unlikely
<i>Opistognathus jacksoniensis</i>	Smiler	Australian. Unlikely
<i>Parablennius yatabei</i>	Blenny	Asia. Not reported in North America. Unlikely
<i>Pagellus bellottii</i>	Red pandora	Eastern Atlantic, Europe and Africa. Unlikely
<i>Paraclinus marmoratus</i>	Marbled blenny	Caribbean, aquarium species. Unlikely
<i>Platyrhina sinensis</i>	Chinese fanray	Asia, not reported in US. Maybe misannotation. Unlikely
<i>Pomadasys sp. CBM:ZF:18112</i>	Grunt	Indian Ocean, Red Sea. Unlikely
<i>Pomadasys stridens</i>	Striped piggy	Indian Ocean, Red Sea. Unlikely
<i>Pomoxis nigromaculatus</i>	Black crappie	Freshwater. Unlikely
<i>Salmo salar</i>	Atlantic salmon	Atlantic, seafood species. Unlikely
<i>Sciaenops ocellatus</i>	Red drum	Atlantic, seafood species. Unlikely
<i>Scorpaena pepo</i>	Pumpkin scorpionfish	Asia, not reported in US only Taiwan. Maybe misannotation. Unlikely
<i>Sphyraena helleri</i>	Heller's barracuda	subtropical, Indopacific. Maybe misannotation. Not reported in North America. Unlikely
<i>Squatina dumeril</i>	Atlantic angel shark	Atlantic coast of US. Maybe misannotation of Pacific Angel Shark. Unlikely
<i>Sternopyx pseudodiaphana</i>	False oblique hatchetfish	Southern hemisphere usually. Maybe misannotation. Unlikely
<i>Synodus intermedius</i>	Sand diver	Subtropical, Caribbean and Atlantic. Maybe misannotation from CA Lizardfish. Unlikely
<i>Synodus saurus</i>	Atlantic lizardfish	Subtropical, Caribbean and Atlantic. Maybe misannotation from CA Lizardfish. Unlikely
<i>Tetratrygon zugei</i>	Pale-edged stingray	Asia, not reported in North America. Unlikely
<i>Tetraodon nigroviridis</i>	Spotted green pufferfish	Freshwater, known aquarium species. Unlikely
<b>Total</b>	<b>44</b>	

A plot of the fish rarefaction curve for eDNA sampling (Figure 3-3) shows the expected levelling-out of diversity estimation with increasing sampling, but does not reach a horizontal plateau, indicating that more eDNA sampling would likely yield a higher estimate of community species richness.

*Figure 3-3. Species rarefaction curve for eDNA sampling of fish species.*

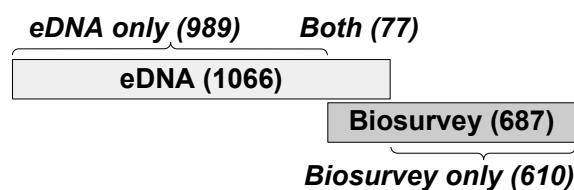


### 3.5 Survey comparisons for invertebrates and algae

The combination of very high diversity and a relatively poor barcode reference library for invertebrates and algae makes eDNA assessments for marine invertebrates more difficult (at present) than for fish. Hence, comparing individual taxa is less useful than for the fish. Because trawls only sample soft-bottom habitat, it is not useful to compare eDNA observations with invertebrate and algae observations from the parallel trawls; a more appropriate comparison is with the unique taxa observed by eDNA and by the 2018 Biosurvey as a whole since additional survey methods were used to characterize invertebrates and algae on nearby hard substrate such as riprap and pier pilings (Figure 3-4).

*Figure 3-4. Counts of invertebrate taxa observed in the eDNA data and Biosurvey.*

*The area where the bars overlap shows species detected by both methods.*



Given the high diversity in the Port area, it is also useful to compare eDNA and Biosurvey data at higher taxonomic levels than species. Figure 3-5 shows counts of genera detected by the two methods.

*Figure 3-5. Counts of invertebrate genera observed in eDNA data and Biosurvey.*

*The area where the bars overlap shows species detected by both methods.*

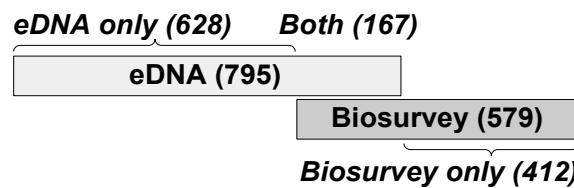
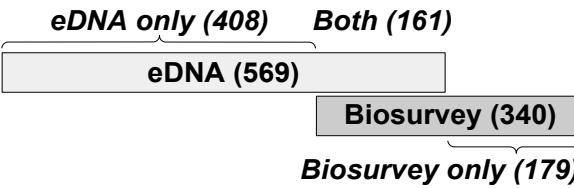


Figure 3-6 shows counts of families detected by the two methods.

*Figure 3-6. Counts of invertebrate families observed in eDNA data and Biosurvey.*

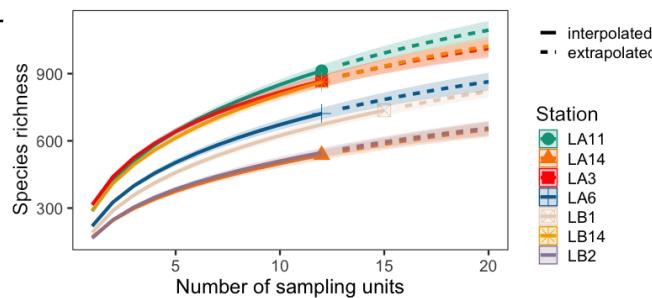
*The area where the bars overlap shows species detected by both methods.*



Grouping taxon observations by higher taxonomic levels leads to a higher proportional overlap between Biosurvey and eDNA detections, but leaves substantial numbers of taxa observed by only one of the two methods.

The rarefaction curve for eDNA sampling for all taxa (Figure 3-7) shows that, as with the fish species rarefaction curve (Figure 3-3), more sampling would likely have yielded a higher number of observed taxa.

*Figure 3-7. Rarefaction curve for eDNA sampling of all invertebrate taxa.*



### 3.6 Detection of known invasive species

Invasive species in the CAL-NEMO database (formerly the NEMESIS and CANOD database) were noted for both the eDNA and Biosurvey observations (Table 3-9). Biosurveys identified a total of 46 of the 53 detected invasive species, while eDNA metabarcoding found a total of 28 invasive species. 25 species were detected only by Biosurveys, 7 species were detected only by eDNA, and 21 species were detected by both methods (Figure 3-8).

*Figure 3-8. Counts of known invasive species observed in eDNA and Biosurvey.*

*The area where the bars overlap shows species detected by both methods.*

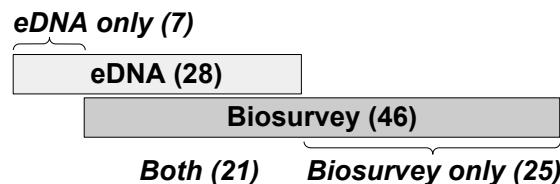


Table 3-9. Invasive species from the Cal-NEMO Database observed in eDNA data Biosurvey.

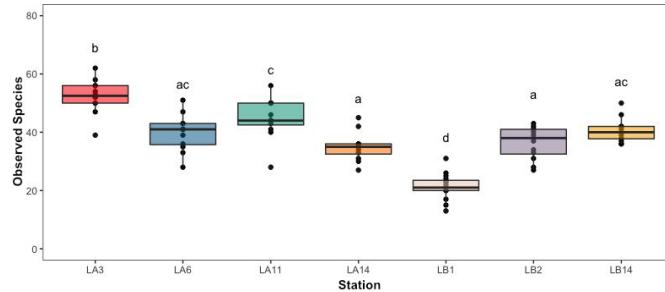
Group	Species	eDNA	Biosurveys	Both
Algae	<i>Lomentaria hakodatensis</i>		✓	
Algae	<i>Sargassum horneri</i>		✓	
Algae	<i>Sargassum muticum</i>		✓	
Algae	<i>Undaria pinnatifida</i>	✓	✓	✓
Annelids - Polychaetes	<i>Branchiomma bairdi</i>		✓	
Annelids - Polychaetes	<i>Hydroides elegans</i>		✓	
Annelids - Polychaetes	<i>Megasyllis nipponica</i>	✓	✓	✓
Annelids - Polychaetes	<i>Myrianida pentadentata</i>	✓		
Annelids - Polychaetes	<i>Polydora cornuta</i>	✓		
Annelids - Polychaetes	<i>Polydora hoplura</i>		✓	
Annelids - Polychaetes	<i>Pseudopolydora paucibranchiata</i>	✓	✓	✓
Bryozoans	<i>Amathia verticillata</i>	✓	✓	✓
Bryozoans	<i>Anguinella palmata</i>	✓	✓	✓
Bryozoans	<i>Bugula neritina</i>	✓		
Bryozoans	<i>Bugulina stolonifera</i>	✓		
Bryozoans	<i>Cryptosula pallasiana</i>	✓	✓	✓
Bryozoans	<i>Watersipora subtorquata complex</i>	✓	✓	✓
Cnidarians - Anthozoans	<i>Nematostella vectensis</i>		✓	
Cnidarians - Hydrozoans	<i>Gonionemus vertens</i>	✓		
Crustaceans - Amphipods	<i>Ampithoe longimana</i>	✓		
Crustaceans - Amphipods	<i>Caprella drepanochir</i>		✓	
Crustaceans - Amphipods	<i>Caprella mutica</i>	✓	✓	✓
Crustaceans - Amphipods	<i>Caprella simia</i>		✓	
Crustaceans - Amphipods	<i>Grandidierella japonica</i>		✓	
Crustaceans - Amphipods	<i>Leucothoe nagatai</i>		✓	
Crustaceans - Amphipods	<i>Monocorophium acherusicum</i>		✓	
Crustaceans - Amphipods	<i>Monocorophium insidiosum</i>		✓	
Crustaceans - Barnacles	<i>Amphibalanus amphitrite</i>	✓	✓	✓
Crustaceans - Isopods	<i>Paranthura japonica</i>		✓	
Crustaceans - Mysids	<i>Deltamysis holmquistae</i>		✓	
Crustaceans - Shrimp	<i>Palaemon macrodactylus</i>		✓	
Fishes	<i>Acanthogobius flavimanus</i>	✓	✓	✓
Fishes	<i>Tridentiger trigonocephalus</i>	✓	✓	✓
Mollusks - Bivalves	<i>Crassostrea gigas</i>		✓	
Mollusks - Bivalves	<i>Musculista senhousia</i>	✓	✓	✓
Mollusks - Bivalves	<i>Mytilus galloprovincialis</i>	✓	✓	✓
Mollusks - Bivalves	<i>Theora lubrica</i>	✓	✓	✓
Mollusks - Bivalves	<i>Venerupis philippinarum</i>		✓	
Mollusks - Gastropods	<i>Crepidula convexa</i>		✓	
Mollusks - Gastropods	<i>Philine auriformis</i>	✓	✓	✓
Tunicates	<i>Ascidia zara</i>	✓	✓	✓
Tunicates	<i>Botrylloides violaceus</i>	✓	✓	✓
Tunicates	<i>Botryllus schlosseri</i>		✓	
Tunicates	<i>Ciona robusta</i>		✓	
Tunicates	<i>Ciona savignyi</i>		✓	
Tunicates	<i>Didemnum vexillum</i>	✓	✓	✓
Tunicates	<i>Diplosoma listerianum</i>	✓		
Tunicates	<i>Microcosmus squamiger</i>	✓	✓	✓
Tunicates	<i>Molgula ficus</i>		✓	
Tunicates	<i>Polyandrocarpa zorritensis</i>		✓	
Tunicates	<i>Styela clava</i>	✓	✓	✓
Tunicates	<i>Styela plicata</i>	✓	✓	✓
Tunicates	<i>Symplegma reptans</i>		✓	
<b>Totals Per Survey</b>		<b>28</b>	<b>46</b>	<b>21</b>
<b>Total Species</b>	<b>53</b>			

### 3.7 Detection of spatial variation by environmental DNA

There are significant differences in the number of fish taxa detected by eDNA across stations and locations. That is, there is evidence of “structure” or non-uniformity across sets of samples taken at different sites and at different locations along transects at sites. We found that the total number of species differed significantly across sites (Figure 3-9), with the highest numbers of species found at LA3, LA11, and LB14 (ANOVA,  $p < 0.001$ ).

*Figure 3-9. Count of fish species observed in eDNA data at each station.*

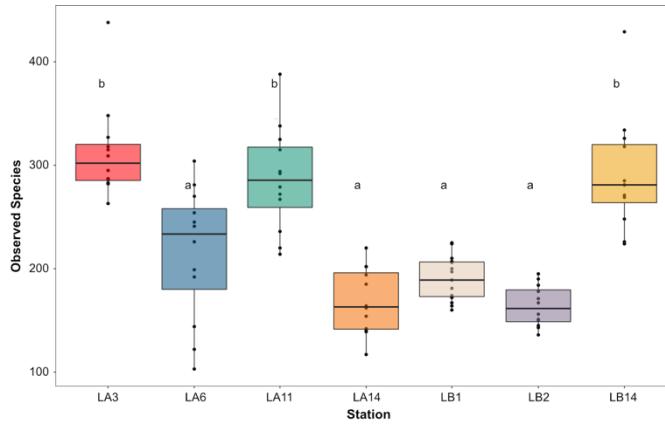
*Dots represent individual samples, boxes denote the interquartile range, and letters indicate groups of stations that do not differ significantly from each other based on*



Similarly, there are significant differences between sites when considering all taxa detected by eDNA (Figure 3-10).

*Figure 3-10. Count of all species observed in eDNA data at each station.*

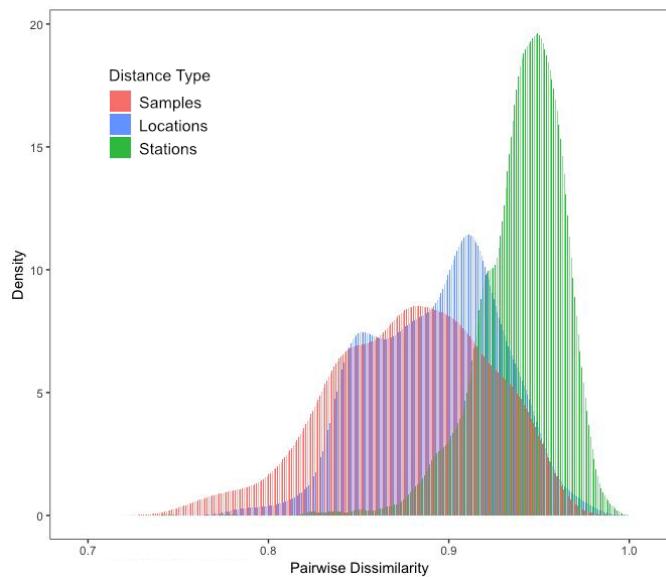
*Dots represent individual samples, boxes denote the interquartile range, and letters indicate groups of stations that do not differ significantly from each other based on ANOVA and Tukey post-hoc tests.*



Bray-Curtis pairwise dissimilarity comparisons found high pairwise dissimilarity across all eDNA samples (Figure 3-11). There is some dissimilarity between replicate samples taken at the same location. However, we found greater dissimilarity between locations along a trawl line than within sample replicates (average distance between the four eDNA sampling locations along a trawl line was about 200 m). Finally, even greater dissimilarity exists between stations across the harbor (see Appendix III for station locations in the harbor). These results are consistent with spatial structuring of eDNA samples, even though there is measurable dissimilarity between replicate samples taken at the same location.

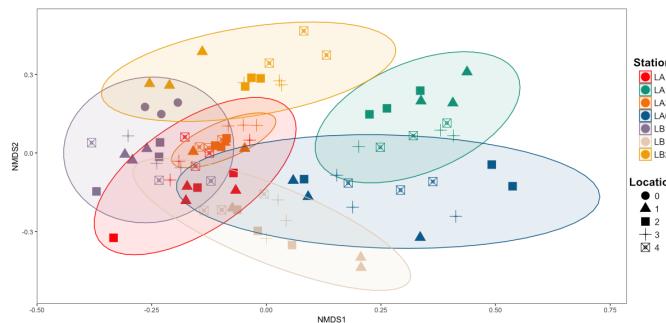
*Figure 3-11. Dissimilarity across biological replicates and sites.*

*Relative distributions of dissimilarity across samples, across locations within trawl tracks, and across stations are shown.*



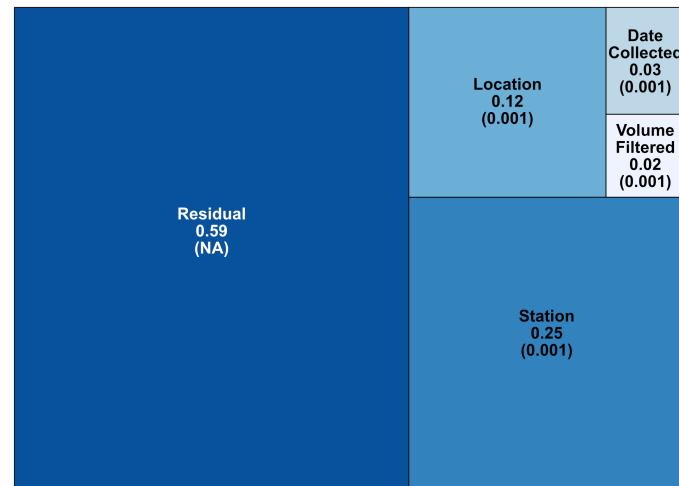
Visualization of the non-metric multidimensional scaling (NMDS) Bray-Curtis dissimilarity ordination (Figure 3-12) showed high overlap between eDNA signatures across the Port. However, despite high overall clustering of samples we found that eDNA signatures differed significantly between sites and between stations (PERMANOVA,  $p=0.001$ , Betadisp  $>0.05$ ).

*Figure 3-12. Bray-Curtis dissimilarity ordination for all taxa.*



Examining the overall apportionment of variance across all sampling factors, there were differences between stations, locations, the two collection dates, and volume of seawater filtered (Figure 3-13). As would be expected, variance contributions from sampling date and volume of seawater filtered are tiny. A residual 59% of the variance remains unexplained by all factors.

Figure 3-13. Apportioned variance for eDNA samples.



### 3.8 Estimation of fish abundance and biomass from environmental DNA

We compared eDNA index scores to abundance of trawled fish and to biomass of trawled fish across the species identified in both survey methods (Figure 3-14 and Figure 3-15). We fit linear models between eDNA index scores and trawled biomass and trawled abundances and calculated  $R^2$  coefficients for each species. Fish relative abundance comparisons between Biosurvey abundance data and eDNA index found high variability in eDNA index scores between stations at each site for all 16 detected species. It is important to note that some of the comparisons that follow, while they may show statistical significance, depend on one to very few individual fish specimens picked up in trawls.

Ten species had a positive correlation between trawl abundance and eDNA index scores (Figure 3-14) with three species displaying a strong positive relationship ( $R^2 > 0.44$ ) and two species displaying a modest positive relationship ( $R^2 > 0.18$ ). All other species had very weak relationships ( $R^2 < 0.1$ ).

Similarly, ten species had a positive correlation between trawl biomass and eDNA index scores (Figure 3-15) with three species displaying a strong positive relationships ( $R^2 > 0.87$ ) and four species displaying a modest positive relationship ( $R^2 > 0.2$ ). All other species had very weak relationships ( $R^2 < 0.1$ ).

Interestingly, eDNA index scores for one species, *Paralichthys californicus*, showed a dramatic shift in relationship between the biomass and abundance comparisons: a much stronger correlation with trawl abundance than with biomass. Three small individuals from LA3 had a combined biomass similar to a single individual from LA6.

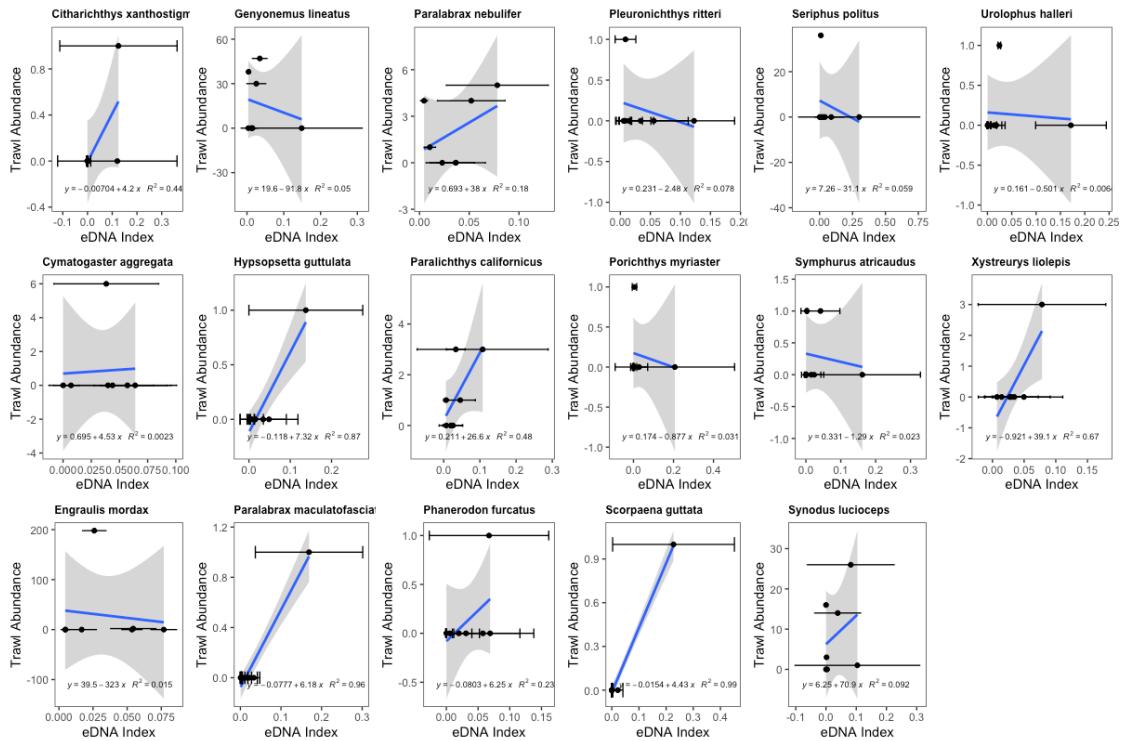


Figure 3-14. eDNA index compared to trawl fish abundance.

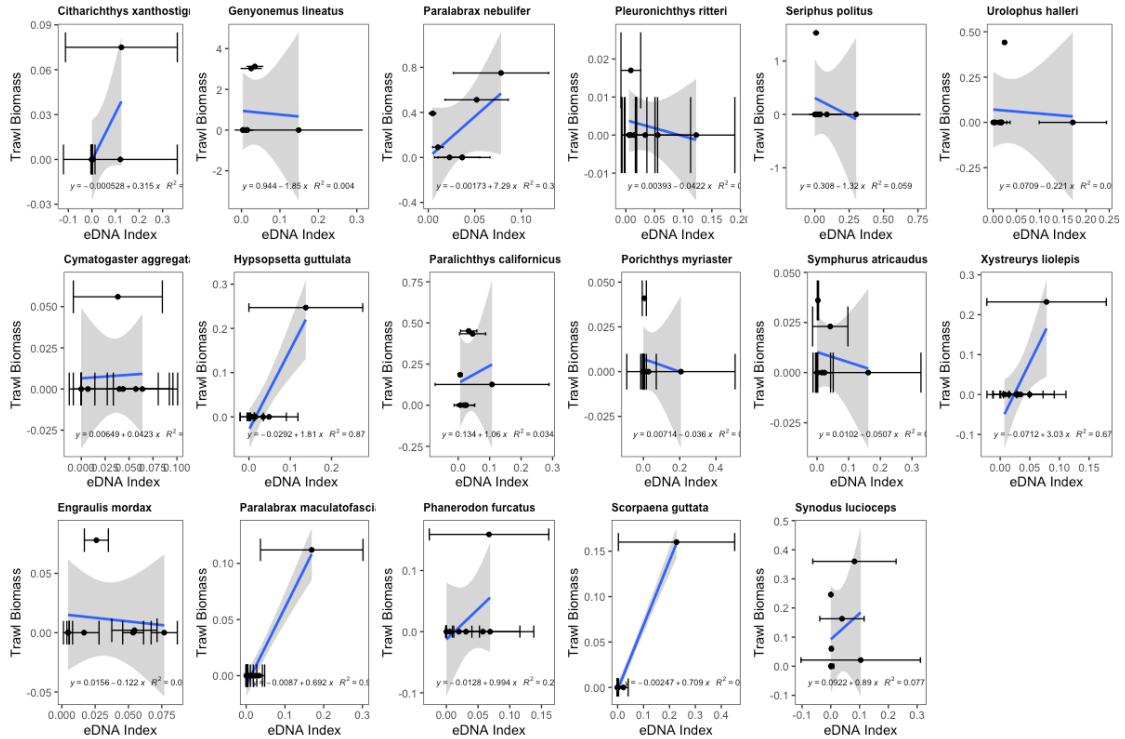


Figure 3-15. eDNA index compared to trawl fish biomass.

---

## 4 DISCUSSION

### 4.1 Comparisons between environmental DNA and conventional sampling

eDNA metabarcoding approaches were successful in detecting a broad range of marine taxa within the Ports of Los Angeles and Long Beach. To date this eDNA survey effort captured the broadest diversity of species in California ever reported (1,195 total marine species) (Port et al. 2016, Andruszkiewicz et al. 2017).

For fish species, it was possible to ensure reference barcode coverage for all taxa reported in the parallel trawl samples by combining publicly-accessible references with references generated specifically for this project (Table 3-3). With that reference coverage for fishes, eDNA recovered all taxa recorded by the trawls (Figure 3-1 and Table 3-4). Additionally, eDNA recovered many (but not all) of the fish taxa reported by all other sampling modalities during the 2018 Biosurvey (Figure 3-2 and Table 3-5) as well as many fish taxa not reported by other modalities (Figure 3-2, Table 3-6, and Table 3-7). This is consistent with previous comparisons of eDNA and conventional surveys (Thomsen et al. 2012a, 2016, Kelly et al. 2017).

Comparisons between methods for invertebrates, involving many more taxa, also show a partial overlap, with eDNA reporting some but not all taxa detected by conventional sampling, and each method reporting taxa unreported by the other (Figure 3-4).

There are several potential reasons for this partial non-overlap of eDNA and conventional ecological methods: (1) lack of complete DNA barcodes available in the reference databases and needed to be able to accurately identify eDNA samples; (2) current limitations in primer design and bioinformatics for DNA barcoding approaches; (3) hyper-sensitivity of eDNA metabarcoding compared to traditional transect surveys; and (4) differences in sampling biases, scope, and scale between eDNA and Biosurvey methods.

#### 4.1.1 Incomplete barcode reference libraries

Species-level resolution of eDNA metabarcoding approaches is constrained by the availability of reference DNA barcodes. Despite California being at the forefront of marine conservation and biotechnology, the vast majority of California's marine species have not been barcoded for the most commonly used eDNA metabarcoding loci (mtDNA 12S, CO1, and 16S genes) (Deiner et al. 2017, Curd et al. 2019). Without matching DNA reference barcodes the analysis cannot assign a taxonomic identity to some sequences. While the 16 reference barcodes for fishes that were created for this study helped to resolve an additional 9 species that were previously not within existing libraries, a lack of tissue from California Scorpionfish (*Scorpaena guttata*) led to the failed species level assignment of eDNA metabarcoding.

There are major efforts underway to barcode the broad diversity of marine life to improve publicly available genetic resources and applied genetic techniques including the CA Fish DNA Sequence

Database, Census of Marine Life, Barcode of Life, and The Diversity Initiative for the Southern California Ocean (DISCO) (Hastings and Burton 2008, Stoeckle and Hebert 2008, Williams et al. 2010, Leray et al. 2013, Wetzer 2015). At present, based on academic efforts, around 41% of fish in Southern California have reference barcodes for the eDNA-relevant 12S genes and 100% of fish in Southern California have reference barcodes for CO1 genes (Hastings and Burton 2008). However, efforts from NOAA, UCLA, and UCSD researchers are on target to complete 12S reference barcodes for an additional 612 species of Southern California fish by the end of 2020 (Gold et al. [submitted]).

For marine invertebrates the reference barcodes are proportionally far less complete. NHMLA's DISCO program has been collecting and registering reference barcodes for marine invertebrates of Southern California for the past three years. About 750 of the estimated 3,000–5,000 described local macroinvertebrate species (defined as larger than a couple of millimeters) are currently registered in BOLD. The remainder need to be collected, identified, vouchered, sequenced, and registered. The DISCO project has been working with numerous agencies and taxonomists to collect molecular-grade specimens and get them reputably identified and registered. Augmenting that reference barcode effort with specimens collected from the Ports biodiversity surveys is the most direct way to ensure the best possible taxonomic coverage for eDNA analysis of Port waters.

Given that reference databases will remain incomplete in the near future, detailed quality assessment and quality control efforts for eDNA-produced species lists are needed to provide important context, particularly in explaining expected taxa that are “missed” by eDNA and non-native taxa whose reference sequence is the best match found for an eDNA sequence. This is especially true if eDNA results are being directly compared to conventional survey methods, as “universal” molecular primers will result in misannotation of species which lack representative reference barcodes. In these instances, eDNA species assignments should always be interpreted within the ecological context of sampling. The majority of eDNA misannotations will be made to sister taxa for which reference barcodes are available, but which may not exist in the surveyed area. Following Occam's razor, it is much more probable that an eDNA read was misannotated than a non-native species is present within a site. Thus adequate efforts should be made to check the accuracy and quality of taxonomic assignments. For sensitive applications of biological monitoring such as those that may trigger violations of local, state, and federal environmental regulations, stricter standards and thresholds should be applied to eDNA identification efforts including the rigorous design and implementation of accurate and reliable specific primers and generation of complete reference databases for taxa of special interest (Kelly et al. 2014, Chambert et al. 2018).

Because of the high diversity of invertebrates and the difficulty of comprehensively resolving species identities over ocean distances, there are certainly numerous taxonomic issues with species names. This is much more of a problem with invertebrates than with fish, just based on the sheer number of taxa. It is very likely that some invertebrate species were actually found in both eDNA and Biosurveys but are listed by different names in the different survey references. Since identifying an overlap depends on having matching names, that will cause an underestimate of the overlap in species observations. Decreasing this problem will only be achieved with more complete reference libraries and accompanying taxonomic and systematic research.

#### *4.1.2 Primer design and bioinformatic limitations*

Primer selection is an important factor of eDNA study design as there is no one “universal” primer that is able to both amplify all species and provide species level resolution across the tree of life (Deiner et al. 2017, Taberlet et al. 2018, Curd et al. 2019). The ability to amplify a broad range of taxa comes at the cost of reduced specificity in taxonomic resolution (Kelly et al. 2019). To achieve the goals of Project Objectives 2 and 3, primers were chosen to target the broadest possible inventory of fish and invertebrate species while achieving the best possible estimate of diversity. The goal was to achieve the best inventory, not to mimic existing sampling methods.

We used four primers known to target broad swaths of marine species (e.g. all bony fishes, 28,000 species or all metazoans +1 million known taxa) (Leray et al. 2013, Miya et al. 2015). The choice of primers with broad taxonomic coverage helps explain the degree of mismatch between eDNA and Biosurvey methods, as eDNA detected a wide variety of species not actively surveyed by the methods used in conventional sampling, including phytoplankton, zooplankton, marine mammals, and terrestrial animals.

An example of compromise between taxonomic breadth and specificity is exemplified by the two 12S primer sets used in this study to target teleosts and elasmobranchs. These primers were able to detect a broad range of taxa spanning a wide diversity of fish life (Miya et al. 2015) However, by designing a primer to amplify all fish taxa with hundreds of millions of years of evolutionary divergence, the primer sacrificed specificity of taxonomic resolution for closely related sister species. In this study, the specificity drawback is seen within rockfishes (genus *Sebastodes*) which are a relatively young adaptive fish radiation (<200,000 years) and vary little within the 12S locus, impacting our the ability to identify *Sebastodes* individuals to species level resolution (Min et al. in prep., Martinez et al. 2017, Yamamoto et al. 2017). One solution to this problem would be to include additional primers. However, each additional primer target essentially doubles laboratory efforts and costs.

Primer design and primer choice should be driven by the research and monitoring goals established at the outset of the study — there is no one right answer. Program choices must evaluate the tradeoffs between taxonomic breadth and specificity, and between cost and generality. As sequencing techniques improve many of these difficult decisions may be alleviated entirely. For example, recent efforts have employed novel hybrid capture arrays to amplify entire mitochondrial genomes (mitogenomes) from eDNA samples, allowing for the capture of dozens of genes needed to make accurate species identifications as well as population genetic inferences (Harper et al. 2019, Sigsgaard et al. 2020).

#### *4.1.3 High sensitivity of eDNA approaches*

This study supports eDNA metabarcoding as a highly sensitive method for marine biodiversity surveying, consistent with prior studies showing greater diversity observed by eDNA than by conventional studies as well as detection of taxa that are difficult to observe (Thomsen et al. 2012b, 2016, Port et al. 2016, Valentini et al. 2016). The high degree of sensitivity is due to the nature of DNA

amplification and its ability to detect fewer than 10 copies of DNA within a single mL, suggesting eDNA metabarcoding can detect one or a few cells (Deiner et al. 2017).

In this pilot study, eDNA results were minimally processed bioinformatically, taking the strategy of reporting any and all sequences that were not eliminated by contamination removal or detection thresholds. This resulted in exposing a large number of observed taxa, contributing to the low degree of overlap between the total number of species between Biosurvey and eDNA methods. eDNA methods identified additional fish species that were not sampled during the 2018 Biosurvey including pelagic and migratory species that are unexpected but not unheard of within or near the Port system such as great white shark (*Carcharodon carcharias*) and thresher shark (*Alopias vulpinis*). eDNA metabarcoding also detected 14 species that were only detected as ichthyoplankton in Biosurveys and not captured as adults, such as the California clingfish (*Gobiesox rhessodon*), shortspine combfish (*Zaniolepis frenata*), longjaw mudsucker (*Gillichthys mirabilis*) and blue-banded goby (*Lythrypnus dalli*).

For targeted investigations, it would be possible to constrain the diversity of results with appropriate primer design or appropriate bioinformatic processing. Taxa that are not of interest to an investigation can always be excluded at any point in the analysis.

Two further examples highlight some of the capabilities of and caveats applicable to eDNA sampling: detection of invasive species, and potential detection of species that are not alive in Port waters.

This study showed that, in the Port environment, Biosurvey methods are capturing the majority of invasive species but eDNA observed an additional 7 invasive species (Table 3-9, but note that definitive determination of the presence of these species will require additional investigation). This is consistent with the finding of Ardura et al. (2015) that eDNA methods can be highly sensitive for detecting a broad range of known serial marine invasives, potentially providing an important tool in a port environment subject to continual visitation by non-local vessels (Zaiko et al. 2015, Rey et al. 2019).

The context of a large and busy port in an urban environment can make it challenging to interpret some eDNA results: the presence of DNA from an organism does not guarantee that it is alive and swimming around nearby. Millions of pounds of seafood are transported in and out of the port by sport and commercial fishermen as well as cargo freight containers (NOAA Fisheries 2019). Inadvertent release of ballast water from cargo ships could potentially introduce DNA to the Port environment even if the organisms are not also discharged or are not living. Additionally, seafood wholesalers ring the Ports, and seafood processing can be a source of measurable eDNA (Yamamoto et al. 2017). eDNA methods are sensitive enough to detect which species of fish was being kept in a cooler when the ice melts into the stormwater system. Runoff from storage and transport of seafood products could lead to the introduction of non-local eDNA into the Ports. Of the fish species that were identified as likely to be non-local in the analysis (Table 3-8) there were 6 species that likely resulted from seafood processing, including Atlantic salmon, red drum, and haddock. The presence of two large aquarium facilities (Aquarium of the Pacific in Long Beach and Cabrillo

Marine Aquarium in San Pedro) are also potential sources of confounding eDNA signatures. Six species of the 44 fish species excluded from the analysis were flagged as likely aquarium species including Cardinalfish (*Apogon imberbis*) and Green Spotted Puffer (*Tetraodon nigroviridis*).

#### 4.1.4 Sampling biases, scope, and scale

All survey methods, including eDNA methods, have their own intrinsic biases. Beach seines, diver surveys, benthic trawls, and bottom grabs yield partially overlapping but complementary biased views of biodiversity, which is why comprehensive surveys use multiple sampling methods. Similarly, we also expect biases with eDNA surveys — it is reasonable to expect eDNA results to differ from those of other methods. Primer selection will affect the species and taxonomic resolution eDNA methods detect. Detections of species using eDNA may arise from non-living organisms or organisms at life stages not captured with traditional methods. Anthropogenic inputs of eDNA can come from sources that do not represent live animals in the area. Spatial and temporal variation can result from eDNA degradation and dispersion rates which are driven by temperature, microbial loads, UV intensity, water chemistry, and water currents (Deiner et al. 2017). These potential sources of biases require consideration to allow appropriate interpretation of results.

Another explanation for the low degree of overlap between Biosurvey and eDNA methods maybe due to the spatial and temporal mismatch between the two methods. The direct comparison between a single trawl and the eDNA from the trawl location yielded little correspondence (Table 3-4). It is reasonable to presume that eDNA samples integrated over a larger temporal and spatial scale than the trawl. Conversely, the multiple-month duration of the Biosurvey undoubtedly integrated over a larger temporal and spatial scale than the two days of eDNA sampling. These differences make it reasonable to expect different results.

This study sampled three distinct 1-liter biological samples (replicates) at four locations along each trawl track (station) to partition eDNA variation along the length of each track. This sampling regime was more intense than typical eDNA sampling efforts which tend to take only 3 replicates at a sampling site. This allowed for investigation of sampling effort and whether the total number of species were saturated at each site. Interestingly, this intensive level of sampling effort was insufficient to capture all taxa within each site as additional replicates failed to saturate, but added additional species (Figure 3-3 and Figure 3-7). These results are similar to those conducted in a controlled mesocosm (aquarium) in which saturation began to occur at 8 replicate 1L samples (Doi et al. 2019). Results of this study suggest that after 10–12 1 L replicates we approach a region of diminishing returns in terms of species accumulation per additional sampling effort. Further work into understanding the biological, technical, spatial, and temporal variability of eDNA signatures on required sampling effort is needed.

## 4.2 Abundance and biomass

An area that clearly warrants further work is eDNA's utility in quantifying the relative abundance or biomass of taxa (Figure 3-14 and Figure 3-15). We found some evidence that eDNA index scores correlate with concurrent trawl biomass and abundance values for some taxa. However, there is

high variability in eDNA index values obtained across biological replicates and samples taken at the same station, and the correlations are not consistent between taxa.

Given our current crude understanding of the origins and fate of DNA in the marine environment, the lack of robustness in this result was to be expected (Taberlet et al. 2012a). However, with better modelling of DNA production and fates, this capability can be expected to improve (Kelly et al. 2019).

We did not quantify the variability in biomass and individuals captured observed between replicate trawl hauls. Previous comparisons of eDNA and trawl or seine replication efforts have demonstrated that both methods have substantial variability (e.g., Shelton et al. 2019). In this study we were unable to investigate the variability of repeated Biosurvey trawls at the same location and thus are unable to quantitatively compare the variability between both methods.

### 4.3 Environmental DNA identifies spatial variation

A key result of this study is that we found high spatial partitioning of eDNA signatures within the Ports of Los Angeles and Long Beach (Figure 3-12 and Figure 3-13). Similar to conventional methods used in the 2018 Biosurvey, eDNA was able to discriminate between different communities of marine species across the Port of LA and Port of Long Beach, allowing for comparisons across regions and monitoring of different Port activities.

Recent work has found that eDNA signatures in the environment can disappear rapidly (<2 hours), and can display spatial structure over less than 5 m (Monuki et al. in prep.). Previous eDNA results have found similar strong spatial distributions of eDNA signatures (Port et al. 2016, Yamamoto et al. 2017, Murakami et al. 2019). The spatial and temporal variation in eDNA signatures are likely controlled by local ocean mixing and microbial community metabolism (Port et al. 2016, Kelly et al. 2018, Murakami et al. 2019). For example, eDNA signatures were found to convey a high degree of spatial separation in a protected nearshore kelp forest ecosystem in Monterey Bay with microhabitat differentiation on the scale of 50 m (Port et al. 2016). In contrast, in more exposed open conditions eDNA signatures were found to vary on scales closer to 800–1000 m (O’Donnell et al. 2017). Together these results strongly suggest that water transport has a controlling effect on the fate and transport of eDNA.

Our results here suggest eDNA signatures varied on spatial scales as smaller than 200 m, similar to the results in a coastal kelp forest (Port et al. 2016). These results suggest the enclosed nature of the Ports, with constrained water flow and circulation relative to open ocean allow for relatively fine scale partitioning of eDNA signatures. Surveys on the outer breakwater of the Ports, more exposed to waves and currents, would be expected to have less spatial resolution (O’Donnell et al. 2017). More work on the spatial and temporal variation of eDNA signatures within marine environments is clearly warranted.

Despite finding evidence of spatial structuring of eDNA samples, site and station accounted for less than 60% of the total variation of eDNA signatures. Previous studies have found high degrees of variability between multiple PCR replicates from the same liter of sea water (termed “technical

replicates”), which may account for a substantial portion of intra-sample variation (>30%) (Port et al. 2016, Kelly et al. 2018). Increasing technical replication increases the overall cost and effort of the sequencing. This study pooled three technical PCR replicates into a single sample to average across this variation at minimum cost. Future studies, especially those focused on detecting invasive or endangered taxa where high confidence of species detections are critical, should employ technical replication in addition to biological replication. That would lead to an increased confidence in eDNA detections and allow for advanced site occupancy statistical modeling to provide confidence scores of species detections and improved interpretation of spatial variation (Schmelzle and Kinziger 2016, Dorazio and Erickson 2018, Doi et al. 2019, Sutter and Kinziger 2019).

---

## 5 PROSPECTS

Every survey method captures a particular aspect of marine biodiversity, and within a highly developed and urbanized environment such as the Ports of Los Angeles and Long Beach, comprehensive surveying is especially challenging. eDNA surveys have the potential to complement conventional sampling, particularly in observing species that are transient or otherwise difficult to observe, in addition to providing a much broader taxonomic scope than is practical to cover in conventional surveys. Conventional methods are still required to capture species once for expert taxonomic determination and deposition in a genetic reference library, but subsequent eDNA monitoring is “takeless”—it inventories biodiversity without impacting it.

Conventional surveys, however, have the ability to capture some types of data that are inaccessible to eDNA. For example, eDNA cannot provide demographic data such as size distribution of fish using the Port Complex. The 2018 Biosurveys have demonstrated that the Ports represent an important nursery habitat for a number of estuarine and offshore fish species (Wood Environment and Infrastructure Solutions, Inc. 2020). Conventional survey methods also allow for additional visual confirmation of important metrics such as sexual maturity in species such as Garibaldi (*Hypsypops rubicundus*), gender ratios in species such as Rock Wrasse (*Halichoeres semicinctus*), and physical abnormalities like those observed in one white croaker during the 2018 Biosurvey (Wood Environment and Infrastructure Solutions, Inc. 2020).

Improving the quality of barcode reference libraries is the single most important contribution for improving eDNA results. This study found higher degrees of overlap in observations between the 2018 Biosurvey and eDNA for fish than for invertebrates. This is unsurprising, given the far better representation of fish taxa than invertebrates in barcode reference databases. After our modest additional fish DNA barcoding effort, eDNA methods still failed to identify 44 fish species identified in the Biosurvey. Of these 44 species, 36 lacked a 12S barcode at the time of this study, which suggests that generating a complete California reference database should be a high priority for improving marine eDNA metabarcoding efforts statewide, beginning with the species identified in the Biosurvey but lacking barcode references. This is even more critical for the less well referenced invertebrates. NHMLA’s DISCO program continues to explore Southern California marine biodiversity and is in the process of expanding reference barcode databases based on extensive historical and new collections. The DISCO project recently (August 2019) held the Los Angeles Urban Ocean Biodiversity Expedition that collected about 4,000 specimens and identified over a thousand taxa (at least 2,000 to species level on initial collection) by world-class taxonomic experts. All these specimens were collected either from inside the Port complex or within 25 miles of it. Continued efforts by DISCO and California marine researchers will improve the capabilities of eDNA methods.

Environmental DNA monitoring in the Port complex promises to expand detection capabilities of endangered, cryptic and invasive species that may evade conventional sampling methods. Tailored monitoring programs for invasive and economically important species utilizing eDNA will be reliant on having a reference barcode for the species of interest. Species of special interest found within the Port complex during the Biosurveys that were notably absent in the eDNA detections

include taxa with commercial value, such as the California spiny lobster, invasive algae *Sargassum horneri* and *Sargassum muticum*, and species with restoration implications such as the white, pink and green abalone. These should be priority species for further barcoding efforts.

It will be important to explore and define the spatial and temporal resolution of eDNA sampling inside and outside the Port environment for effective planning of eDNA inventory projects. This study demonstrated spatial variability in eDNA samples between areas of the Port, but could not encompass the experimental design to quantitatively characterize spatial and temporal resolution. That determination will require a series of samples through time at a selection of sites, sampled and analyzed specifically to define those parameters. Knowing how eDNA signals vary across time and space will allow the design of efficient sampling design for eDNA inventories that will be able to accurately document differences across the Port complex and changes through time in response to environmental remediation.

Interpretation of eDNA results must be done with an awareness of the high sensitivity of eDNA. Although this method observes species which would be difficult to detect with conventional sampling, it can also observe species whose DNA may be present from confounding sources such as seafood processing, aquarium proximity, and urban influence.

In the future, it may be possible to use eDNA-derived metrics such as the eDNA index to give relative estimates of the abundance of particular taxa in different areas or at different times. Recent studies, mostly focussing on single species, have reported encouraging results when carefully designed to measure biomass or number of individuals (Nevers et al. 2018, Kelly et al. 2019, Akamatsu et al. 2020). Given this is a nascent and rapidly developing tool, it is not reasonable to expect broadly applicable biomass or number estimates from eDNA. However, species-specific quantity estimates seem well on their way to becoming a potential survey tool as laboratory techniques improve and sequencing costs decline (Ushio et al. 2017).

With conventional biological surveys, the project report provides the snapshot of the biodiversity evaluation at a given time, and that summary is usually final. Though some studies may voucher every specimen captured for later re-evaluation, it is expensive to maintain voucher collections in the long term, and expensive to reassess species determinations. In contrast, sequence datasets from eDNA are inexpensive to maintain and can be reanalyzed digitally as barcode reference libraries grow, improving taxonomic resolution. For example, the current study results could be re-analyzed when the NHMLA completes its barcoding of the L.A. Urban Ocean Expedition taxa, greatly improving the resolution of invertebrate taxa. The eDNA sequences recorded from the eDNA sampling of this study are an important genetic resource whose ability to provide information will improve with time.

Additionally, the original eDNA samples from this study are stored permanently in the CALeDNA sample repository (Meyer et al. 2019) at a cost that is trivial compared with storing permanent specimen vouchers. As novel genetic analysis technologies are developed in the future, these samples can be reanalyzed.

Sequence datasets whose resolution improves with time and readily available physical samples that can be reanalyzed provide a time-machine-like capability for the future. It will be possible to reach back in time and enhance the results from this dataset for improved before-and-after studies that document change through time. This is a valuable resource for future managers, conservation biologists, and biodiversity scientists. In this way, eDNA strongly augments results from conventional surveys.

---

## 6 REFERENCES CITED

- Akamatsu Y, Kume G, Gotou M, Kono T, Fujii T, Inui R, Kurita Y (2020) Using environmental DNA analyses to assess the occurrence and abundance of the endangered amphidromous fish *Plecoglossus altivelis ryukyuensis*. *Biodiversity Data Journal* 8: e39679. <https://doi.org/10.3897/BDJ.8.e39679>
- Amir A, McDonald D, Navas-Molina JA, Kopylova E, Morton JT, Xu ZZ, Kightley EP, Thompson LR, Hyde ER, Gonzalez A (2017) Deblur rapidly resolves single-nucleotide community sequence patterns. *mSystems* 2: e00191–16. <https://doi.org/10.1128/mSystems.00191-16>
- Andruszkiewicz EA, Starks HA, Chavez FP, Sassoubre LM, Block BA, Boehm AB (2017) Biomonitoring of marine vertebrates in Monterey Bay using eDNA metabarcoding. Doi H (Ed). *PLoS ONE* 12: e0176343. <https://doi.org/10.1371/journal.pone.0176343>
- Appeltans W, Ahyong ST, Anderson G, Angel MV, Artois T, Bailly N, Bamber R, Barber A, Bartsch I, Berta A, Błażewicz-Paszkowycz M, Bock P, Boxshall G, Boyko CB, Brandão SN, Bray RA, Bruce NL, Cairns SD, Chan T-Y, Cheng L, Collins AG, Cribb T, Curini-Galletti M, Dahdouh-Guebas F, Davie PJF, Dawson MN, De Clerck O, Decock W, De Grave S, de Voogd NJ, Domning DP, Emig CC, Erséus C, Eschmeyer W, Fauchald K, Fautin DG, Feist SW, Fransen CHJM, Furuya H, Garcia-Alvarez O, Gerken S, Gibson D, Gittenberger A, Gofas S, Gómez-Daglio L, Gordon DP, Guiry MD, Hernandez F, Hoeksema BW, Hopcroft RR, Jaume D, Kirk P, Koedam N, Koenemann S, Kolb JB, Kristensen RM, Kroh A, Lambert G, Lazarus DB, Lemaitre R, Longshaw M, Lowry J, Macpherson E, Madin LP, Mah C, Mapstone G, McLaughlin PA, Mees J, Meland K, Messing CG, Mills CE, Molodtsova TN, Mooi R, Neuhaus B, Ng PKL, Nielsen C, Norenburg J, Opresko DM, Osawa M, Paulay G, Perrin W, Pilger JF, Poore GCB, Pugh P, Read GB, Reimer JD, Rius M, Rocha RM, Saiz-Salinas JI, Scarabino V, Schierwater B, Schmidt-Rhaesa A, Schnabel KE, Schotte M, Schuchert P, Schwabe E, Segers H, Self-Sullivan C, Shenkar N, Siegel V, Sterrer W, Stöhr S, Swalla B, Tasker ML, Thuesen EV, Timm T, Todaro MA, Turon X, Tyler S, Uetz P, van der Land J, Vanhoorne B, van Ofwegen LP, van Soest RWM, Vanaverbeke J, Walker-Smith G, Walter TC, Warren A, Williams GC, Wilson SP, Costello MJ (2012) The magnitude of global marine species diversity. *Current Biology* 22: 2189–2202. <https://doi.org/10.1016/j.cub.2012.09.036>
- Ardura A, Zaiko A, Martinez JL, Samuloviene A, Borrell Y, Garcia-Vazquez E (2015) Environmental DNA evidence of transfer of North Sea molluscs across tropical waters through ballast water. *Journal of Molluscan Studies* 81: 495–501. <https://doi.org/10.1093/mollus/eyv022>
- Barnes MA, Turner CR (2016) The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics* 17: 1–17. <https://doi.org/10.1007/s10592-015-0775-4>
- Barnes MA, Turner CR, Jerde CL, Renshaw MA, Chadderton WL, Lodge DM (2014) Environmental conditions influence eDNA persistence in aquatic systems. *Environmental Science & Technology* 48: 1819–1827. <https://doi.org/10.1021/es404734p>
- Beheregaray LB, Caccone A (2007) Cryptic biodiversity in a changing world. *Journal of Biology* 6: 9. <https://doi.org/10.1186/jbiol60>
- Beja-Pereira A, Oliveira R, Alves PC, Schwartz MK, Luikart G (2009) Advancing ecological understandings through technological transformations in noninvasive genetics. *Molecular Ecology Resources* 9: 1279–1301. <https://doi.org/10.1111/j.1755-0998.2009.02699.x>

- Bohmann K, Evans A, Gilbert MTP, Carvalho GR, Creer S, Knapp M, Yu DW, de Bruyn M (2014) Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution* 29: 358–367. <https://doi.org/10.1016/j.tree.2014.04.003>
- Boussarie G, Bakker J, Wangensteen OS, Mariani S, Bonnin L, Juhel J-B, Kiszka JJ, Kulbicki M, Manel S, Robbins WD, Vigliola L, Mouillot D (2018) Environmental DNA illuminates the dark diversity of sharks. *Science Advances* 4: eaap9661. <https://doi.org/10.1126/sci-adv.aap9661>
- Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: high-resolution sample inference from Illumina amplicon data. *Nature methods* 13: 581. <https://doi.org/10.1038/nmeth.3869>
- Chambert T, Pilliod DS, Goldberg CS, Doi H, Takahara T (2018) An analytical framework for estimating aquatic species density from environmental DNA. *Ecology and Evolution* 8: 3468–3477. <https://doi.org/10.1002/ece3.3764>
- Chao A, Gotelli NJ, Hsieh TC, Sande EL, Ma KH, Colwell RK, Ellison AM (2014) Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecological Monographs* 84: 45–67. <https://doi.org/10.1890/13-0133.1>
- Collins RA, Wangensteen OS, O’Gorman EJ, Mariani S, Sims DW, Genner MJ (2018) Persistence of environmental DNA in marine systems. *Communications Biology* 1: 185. <https://doi.org/10.1038/s42003-018-0192-6>
- Costello M, Fleharty M, Abreu J, Farjoun Y, Ferriera S, Holmes L, Granger B, Green L, Howd T, Mason T (2018) Characterization and remediation of sample index swaps by non-redundant dual indexing on massively parallel sequencing platforms. *BMC genomics* 19: 332. <https://doi.org/10.1186/s12864-018-4703-0>
- Curd EE, Gold Z, Kandlikar GS, Gomer J, Ogden M, O’Connell T, Pipes L, Schweizer TM, Rabichow L, Lin M, Shi B, Barber PH, Kraft N, Wayne R, Meyer RS (2019) *Anacapa Toolkit*: an environmental DNA toolkit for processing multilocus metabarcode datasets. *Methods in Ecology and Evolution*. <https://doi.org/10.1111/2041-210X.13214>
- Deiner K, Bik HM, Mächler E, Seymour M, Lacoursière-Roussel A, Altermatt F, Creer S, Bista I, Lodge DM, de Vere N, Pfrender ME, Bernatchez L (2017) Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology* 26: 5872–5895. <https://doi.org/10.1111/mec.14350>
- Dejean T, Valentini A, Duparc A, Pellier-Cuit S, Pompanon F, Taberlet P, Miaud C (2011) Persistence of environmental DNA in freshwater ecosystems. Gilbert JA (Ed). *PLoS ONE* 6: e23398. <https://doi.org/10.1371/journal.pone.0023398>
- Doi H, Fukaya K, Oka S, Sato K, Kondoh M, Miya M (2019) Evaluation of detection probabilities at the water-filtering and initial PCR steps in environmental DNA metabarcoding using a multispecies site occupancy model. *Scientific Reports* 9: 3581. <https://doi.org/10.1038/s41598-019-40233-1>
- Dorazio RM, Erickson RA (2018) EDNAOCCUPANCY: An R package for multiscale occupancy modeling of environmental DNA data. *Molecular Ecology Resources* 18: 368–380. <https://doi.org/10.1111/1755-0998.12735>
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792–1797. <https://doi.org/10.1093/nar/gkh340>

- Evans NT, Olds BP, Renshaw MA, Turner CR, Li Y, Jerde CL, Mahon AR, Pfrender ME, Lamberti GA, Lodge DM (2016) Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. *Molecular Ecology Resources* 16: 29–41. <https://doi.org/10.1111/1755-0998.12433>
- Faircloth BC, Glenn TC (2014) Protocol: Preparation of an Ampure XP substitute (aka Serapure). Available from: [http://baddna.uga.edu/Protocols/Speedbead\\_Protocol\\_June2016.docx](http://baddna.uga.edu/Protocols/Speedbead_Protocol_June2016.docx).
- Fofonoff PW, Ruiz GM, Steves B, Simkanin C, Carlton JT (2019) California Non-native Estuarine and Marine Organisms (Cal-NEMO) System. Available from: <http://invasions.si.edu/nemesis/> (October 27, 2019).
- Froese R, Pauly D (2019) FishBase World Wide Web electronic publication. FishBase. Available from: <http://www.fishbase.org> (May 26, 2020).
- Gallego R (2019) Ramon Gallego — GitHub repository. GitHub. Available from: <https://github.com/ramongallego?tab=repositories> (October 10, 2019).
- Gao X, Lin H, Revanna K, Dong Q (2017) A Bayesian taxonomic classification method for 16S rRNA gene sequences with improved species-level accuracy. *BMC Bioinformatics* 18: 247. <https://doi.org/10.1186/s12859-017-1670-4>
- 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: 851–861. <https://doi.org/10.1111/1755-0998.12138>
- Gold ZJ (2020) Design and implementation of environmental DNA metabarcoding methods for monitoring the Southern California Marine Protected Area Network. Ph. D. Thesis. University of California, Los Angeles Available from: <https://escholarship.org/uc/item/7zf5g0cb>.
- Gold ZJ, Choi E, Kacev D, et al. ([submitted]) FishCARD: Fish 12S California Current specific reference database for enhanced metabarcoding efforts. *Molecular Ecology Resources*.
- Goldberg CS, Turner CR, Deiner K, Klymus KE, Thomsen PF, Murphy MA, Spear SF, McKee A, Oyler-McCance SJ, Cornman RS, Laramie MB, Mahon AR, Lance RF, Pilliod DS, Strickler KM, Waits LP, Fremier AK, Takahara T, Herder JE, Taberlet P (2016) Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution* 7: 1299–1307. <https://doi.org/10.1111/2041-210X.12595>
- Harper LR, Lawson Handley L, Carpenter AI, Ghazali M, Di Muri C, Macgregor CJ, Logan TW, Law A, Breithaupt T, Read DS, McDevitt AD, Häneling B (2019) Environmental DNA (eDNA) metabarcoding of pond water as a tool to survey conservation and management priority mammals. *Biological Conservation* 238: 108225. <https://doi.org/10.1016/j.biocon.2019.108225>
- Hastings PA, Burton RS (2008) Establishing a DNA sequence database for the marine fish fauna of California. University of California, San Diego, 4pp. Available from: <https://escholarship.org/content/qt1ck9b3qs/qt1ck9b3qs.pdf>.
- Heck KL Jr, van Belle G, Simberloff D (1975) Explicit calculation of the rarefaction diversity measurement and the determination of sufficient sample size. *Ecology* 56: 1459–1461. <https://doi.org/10.2307/1934716>
- Hsieh TC, Ma KH, Chao A (2019) iNEXT: Interpolation and Extrapolation for Species Diversity. Available from: <https://cran.r-project.org/web/packages/iNEXT/index.html>.

- Kandlikar GS, Gold ZJ, Cowen MC, Meyer RS, Freise AC, Kraft NJ, Moberg-Parker J, Sprague J, Kushner DJ, Curd EE (2018) ranacapa: An R package and Shiny web app to explore environmental DNA data with exploratory statistics and interactive visualizations. F1000Research 7. <https://doi.org/10.12688/f1000research.16680.1>
- Karanovic T, Djurakic M, Eberhard SM (2016) Cryptic species or inadequate taxonomy? Implementation of 2D geometric morphometrics based on integumental organs as landmarks for delimitation and description of copepod taxa. Systematic Biology 65: 304–327. <https://doi.org/10.1093/sysbio/syv088>
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28: 1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>
- Kelly RP (2016) Making environmental DNA count. Molecular Ecology Resources 16: 10–12. <https://doi.org/10.1111/1755-0998.12455>
- Kelly RP, Gallego R, Jacobs-Palmer E (2018) The effect of tides on nearshore environmental DNA. PeerJ 6: e4521. <https://doi.org/10.7717/peerj.4521>
- Kelly RP, Shelton AO, Gallego R (2019) Understanding PCR processes to draw meaningful conclusions from environmental DNA studies. Scientific Reports 9: 1–14. <https://doi.org/10.1038/s41598-019-48546-x>
- Kelly RP, Closek CJ, O'Donnell JL, Kralj JE, Shelton AO, Samhouri JF (2017) Genetic and manual survey methods yield different and complementary views of an ecosystem. Frontiers in Marine Science 3: 283. <https://doi.org/10.3389/fmars.2016.00283>
- Kelly RP, O'Donnell JL, Lowell NC, Shelton AO, Samhouri JF, Hennessey SM, Feist BE, Williams GD (2016) Genetic signatures of ecological diversity along an urbanization gradient. PeerJ 4: e2444. <https://doi.org/10.7717/peerj.2444>
- Kelly RP, Port JA, Yamahara KM, Martone RG, Lowell N, Thomsen PF, Mach ME, Bennett M, Prahler E, Caldwell MR, Crowder LB (2014) Harnessing DNA to improve environmental management. Science 344: 1455–1456. <https://doi.org/10.1126/science.1251156>
- Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nature Methods 9: 357–359. <https://doi.org/10.1038/nmeth.1923>
- Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Boehm JT, Machida RJ (2013) A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. Frontiers in Zoology 10: 34. <https://doi.org/10.1186/1742-9994-10-34>
- Lewin HA, Robinson GE, Kress WJ, Baker WJ, Coddington J, Crandall KA, Durbin R, Edwards SV, Forest F, Gilbert MTP, Goldstein MM, Grigoriev IV, Hackett KJ, Haussler D, Jarvis ED, Johnson WE, Patrinos A, Richards S, Castilla-Rubio JC, van Sluys M-A, Soltis PS, Xu X, Yang H, Zhang G (2018) Earth BioGenome Project: Sequencing life for the future of life. Proceedings of the National Academy of Sciences 115: 4325–4333. <https://doi.org/10.1073/pnas.1720115115>
- Machida RJ, Leray M, Ho S-L, Knowlton N (2017) Metazoan mitochondrial gene sequence reference datasets for taxonomic assignment of environmental samples. Scientific Data 4: 170027. <https://doi.org/10.1038/sdata.2017.27>

- Martinez E, Buonaccorsi V, Hyde JR, Aguilar A (2017) Population genomics reveals high gene flow in grass rockfish (*Sebastodes rastrelliger*). *Marine genomics* 33: 57–63.  
<https://doi.org/10.1016/j.margen.2017.01.004>
- McMurdie PJ, Holmes S (2013) phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* 8: e61217. <https://doi.org/10.1371/journal.pone.0061217>
- Meyer R, Curd EE, Schweizer T, Gold Z, Ramos DR, Shirazi S, Kandlikar G, Kwan W-Y, Lin M, Friese A (2019) The California Environmental DNA “CALeDNA” program. *bioRxiv*.  
<https://doi.org/10.1101/503383>
- Min M, Barber PH, Gold ZJ (in prep.) MiSebastes: An eDNA metabarcoding primer for rockfishes (*Sebastodes*).
- Miya M, Sato Y, Fukunaga T, Sado T, Poulsen JY, Sato K, Minamoto T, Yamamoto S, Yamanaka H, Araki H, Kondoh M, Iwasaki W (2015) 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: 150088. <https://doi.org/10.1098/rsos.150088>
- Monuki KS, Barber PH, Gold ZJ (in prep.) Fine scale depth variation of eDNA signatures: informing nearshore coastal marine protected area monitoring.
- Murakami H, Yoon S, Kasai A, Minamoto T, Yamamoto S, Sakata MK, Horiuchi T, Sawada H, Kondoh M, Yamashita Y, Masuda R (2019) Dispersion and degradation of environmental DNA from caged fish in a marine environment. *Fisheries Science* 85: 327–337.  
<https://doi.org/10.1007/s12562-018-1282-6>
- Nevers MB, Byappanahalli MN, Morris CC, Shively D, Przybyla-Kelly K, Spoljaric AM, Dickey J, Roseman EF (2018) Environmental DNA (eDNA): A tool for quantifying the abundant but elusive round goby (*Neogobius melanostomus*). Doi H (Ed). *PLoS ONE* 13: e0191720.  
<https://doi.org/10.1371/journal.pone.0191720>
- NOAA Fisheries (2019) Foreign fishery trade data. NOAA Fisheries. Available from:  
<https://www.fisheries.noaa.gov/national/sustainable-fisheries/foreign-fishery-trade-data> (October 11, 2019).
- O'Donnell JL, Kelly RP, Shelton AO, Samhouri JF, Lowell NC, Williams GD (2017) Spatial distribution of environmental DNA in a nearshore marine habitat. *PeerJ* 5: e3044.  
<https://doi.org/10.7717/peerj.3044>
- Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2019) vegan: Community Ecology Package. Available from: <https://CRAN.R-project.org/package=vegan>.
- Pérez-Ponce de León G, Poulin R (2016) Taxonomic distribution of cryptic diversity among metazoans: not so homogeneous after all. *Biology Letters* 12: 20160371.  
<https://doi.org/10.1098/rsbl.2016.0371>
- Pfenninger M, Schwenk K (2007) Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. *BMC Evolutionary Biology* 7: 121.  
<https://doi.org/10.1186/1471-2148-7-121>
- Port JA, O'Donnell JL, Romero-Maraccini OC, Leary PR, Litvin SY, Nickols KJ, Yamahara KM, Kelly RP (2016) Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. *Molecular Ecology* 25: 527–541. <https://doi.org/10.1111/mec.13481>

- R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from: <https://www.R-project.org/>.
- Ratnasingham S, Hebert PDN (2007) BOLD: The Barcode of Life Data System (<http://www.barcodinglife.org>). Molecular Ecology Notes 7: 355–364. <https://doi.org/10.1111/j.1471-8286.2006.01678.x>
- Rey A, Carney KJ, Quinones LE, Pagenkopp Lohan KM, Ruiz GM, Basurko OC, Rodríguez-Ezpeleta N (2019) Environmental DNA metabarcoding: a promising tool for ballast water monitoring. Environmental science & technology 53: 11849–11859. <https://doi.org/10.1021/acs.est.9b01855>
- Sassoubre LM, Yamahara KM, Gardner LD, Block BA, Boehm AB (2016) Quantification of environmental DNA (eDNA) shedding and decay rates for three marine fish. Environmental Science & Technology 50: 10456–10464. <https://doi.org/10.1021/acs.est.6b03114>
- Schmelzle MC, Kinziger AP (2016) Using occupancy modelling to compare environmental DNA to traditional field methods for regional-scale monitoring of an endangered aquatic species. Molecular Ecology Resources 16: 895–908. <https://doi.org/10.1111/1755-0998.12501>
- Shelton AO, Kelly RP, O'Donnell JL, Park L, Schwenke P, Greene C, Henderson RA, Beamer EM (2019) Environmental DNA provides quantitative estimates of a threatened salmon species. Biological Conservation 237: 383–391. <https://doi.org/10.1016/j.biocon.2019.07.003>
- Sigsgaard EE, Jensen MR, Winkelmann IE, Møller PR, Hansen MM, Thomsen PF (2020) Population-level inferences from environmental DNA—Current status and future perspectives. Evolutionary Applications 13: 245–262. <https://doi.org/10.1111/eva.12882>
- Spens J, Evans AR, Halfmaerten D, Knudsen SW, Sengupta ME, Mak SST, Sigsgaard EE, Hellström M (2017) Comparison of capture and storage methods for aqueous microbial eDNA using an optimized extraction protocol: advantage of enclosed filter. Methods in Ecology and Evolution 8: 635–645. <https://doi.org/10.1111/2041-210X.12683>
- Stoeckle MY, Hebert PDN (2008) Barcode of life: DNA tags help classify animals. Scientific American 299: 82–89. <https://doi.org/10.1038/scientificamerican1008-82>
- Sutter M, Kinziger AP (2019) Rangewide tidewater goby occupancy survey using environmental DNA. Conservation Genetics. <https://doi.org/10.1007/s10592-019-01161-9>
- Taberlet P, Coissac E, Hajibabaei M, Rieseberg LH (2012a) Environmental DNA. Molecular Ecology 21: 1789–1793. <https://doi.org/10.1111/j.1365-294X.2012.05542.x>
- Taberlet P, Bonin A, Zinger L, Coissac E (2018) Environmental DNA: for biodiversity research and monitoring. Oxford University Press.
- Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012b) Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular Ecology 21: 2045–2050. <https://doi.org/10.1111/j.1365-294X.2012.05470.x>
- Thomsen PF, Kielgast J, Iversen LL, Møller PR, Rasmussen M, Willerslev E (2012a) Detection of a diverse marine fish fauna using environmental DNA from seawater samples. Lin S (Ed). PLoS ONE 7: e41732. <https://doi.org/10.1371/journal.pone.0041732>
- Thomsen PF, Møller PR, Sigsgaard EE, Knudsen SW, Jørgensen OA, Willerslev E (2016) Environmental DNA from seawater samples correlate with trawl catches of subarctic, deepwater

fishes. Mahon AR (Ed). PLoS ONE 11: e0165252. <https://doi.org/10.1371/journal.pone.0165252>

Thomsen PF, Kielgast J, Iversen LL, Wiuf C, Rasmussen M, Gilbert MTP, Orlando L, Willerslev E (2012b) Monitoring endangered freshwater biodiversity using environmental DNA. Molecular Ecology 21: 2565–2573. <https://doi.org/10.1111/j.1365-294X.2011.05418.x>

Tillotson MD, Kelly RP, Duda JJ, Hoy M, Kralj J, Quinn TP (2018) Concentrations of environmental DNA (eDNA) reflect spawning salmon abundance at fine spatial and temporal scales. Biological Conservation 220: 1–11. <https://doi.org/10.1016/j.biocon.2018.01.030>

Ushio M, Murakami H, Masuda R, Sado T, Miya M, Sakurai S, Yamanaka H, Minamoto T, Kondoh M (2017) Quantitative monitoring of multispecies fish environmental DNA using high-throughput sequencing. Metabarcoding and Metagenomics 2: 1–15. <https://doi.org/10.1101/113472>

Valentini A, Taberlet P, Miaud C, Civade R, Herder J, Thomsen PF, Bellemain E, Besnard A, Coissac E, Boyer F, Gaboriaud C, Jean P, Poulet N, Roset N, Copp GH, Geniez P, Pont D, Argillier C, Baudoin J-M, Peroux T, Crivelli AJ, Olivier A, Acqueberge M, Le Brun M, Møller PR, Willerslev E, Dejean T (2016) Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. Molecular Ecology 25: 929–942. <https://doi.org/10.1111/mec.13428>

Wetzer R (2015) Collecting and preserving marine and freshwater Isopoda (Crustacea: Peracarida). Biodiversity Data Journal 3: e4912. <https://doi.org/10.3897/BDJ.3.e4912>

Wilkins D (2019) treemapify: Draw Treemaps in “ggplot2.” Available from: <https://CRAN.R-project.org/package=treemapify>.

Williams MJ, Ausubel J, Painer I, Garcia SM, Baker DJ, Clark MR, Mannix H, Yarincik K, Halpin PN (2010) Making marine life count: a new baseline for policy. PLoS biology 8. <https://doi.org/10.1371/journal.pbio.1000531>

Wood Environment and Infrastructure Solutions, Inc. (2020) 2018 Biological Surveys of the Los Angles and Long Beach Harbors. Prepared for the Port of Los Angeles and Long Beach

Yamamoto S, Masuda R, Sato Y, Sado T, Araki H, Kondoh M, Minamoto T, Miya M (2017) Environmental DNA metabarcoding reveals local fish communities in a species-rich coastal sea. Scientific Reports 7: 40368. <https://doi.org/10.1038/srep40368>

Zaiko A, Martinez JL, Ardura A, Clusa L, Borrell YJ, Samuiloviene A, Roca A, Garcia-Vazquez E (2015) Detecting nuisance species using NGST: methodology shortcomings and possible application in ballast water monitoring. Marine Environmental Research 112: 64–72. <https://doi.org/10.1016/j.marenvres.2015.07.002>

---

## APPENDIX I. INVERTEBRATE REFERENCE BARCODES

Benthic marine invertebrates for which reference barcodes sequences (298) were generated for this project. Initial number gives the count of sequences generated for each species (multiple specimens are required to define the genetic variance within a nominal species).

- 2 Actiniidae: *Anthopleura elegantissima*
- 1 Actiniidae: *Epiactis prolifera*
- 2 Aegiridae: *Aegires albopunctatus*
- 1 Aeolidiidae: *Aeolidia papillosa*
- 1 Aglajidae: *Navanax inermis*
- 1 Aglaopheniidae: *Aglaophenia*
- 1 Alpheidae: *Alpheus bellimanus*
- 1 Ammotheidae: *Ammothaea hilgendorfi*
- 1 Amphiuridae: *Amphiodia occidentalis*
- 2 Ampithoidae: *Ampithoe plumulosa*
- 1 Aplysiidae: *Aplysia californica*
- 1 Arminidae: *Armina californica*
- 1 Ascorhynchidae: *Eurycyde spinosa*
- 3 Asterinidae: *Patiria miniata*
- 1 Astropectinidae: *Astropecten armatus*
- 2 Astropectinidae: *Astropecten californicus*
- 2 Balanidae: *Balanus glandula*
- 2 Balanidae: *Megabalanus*
- 1 Buccinidae: *Kelletia kelletii*
- 1 Bugulidae: *Bugula neritina*
- 2 Bursidae: *Crossata californica*
- 2 Cadiinidae: *Cadlina flavomaculata*
- 2 Caligidae: *Caligus klawai*
- 1 Callistoplacidae: *Callistochiton decoratus*
- 1 Calyptraeidae: *Crepidula naticarum*
- 2 Calyptraeidae: *Crepidatella lingulata*
- 3 Cancridae: *Metacarcinus anthonyi*
- 1 Cancridae: *Romaleon jordani*
- 2 Caprellidae: *Caprella californica*
- 1 Chaetopteridae: *Phyllochaetoperus socialis*
- 1 Chevaliidiae: *Chevalia inaequalis*
- 1 Chromodorididae: *Felimida macfarlandi*
- 2 Chrysopetalidae: *Paleanotus bellis*
- 1 Cionidae: *Ciona robusta*
- 2 Cionidae: *Ciona savignyi*
- 5 Columbellidae: *Alia carinata*
- 2 Conidae: *Californiconus californicus*
- 3 Corambidae: *Corambe steinbergae*
- 2 Corophiidae: *Laticorophium baconi*
- 4 Crangonidae: *Crangon nigromaculata*
- 1 Crangonidae: *Metacrangon spinosissima*
- 1 Cucumariidae: *Cucumaria miniata*
- 2 Cypraeidae: *Neobernaya spadicea*
- 2 Dendrodorididae: *Doriopsilla albopunctata*

- 1 Dendrodorididae: *Doriopsilla fulva*  
2 Dendronotidae: *Dendronotus iris*  
2 Dendronotidae: *Dendronotus venustus*  
1 Dentrastidae: *Dendraster excentricus*  
2 Discodorididae: *Diaulula sandiegensis*  
1 Discodorididae: *Rostanga pulchra*  
1 Epialtidae: *Epialtoides hiltoni*  
1 Epialtidae: *Loxorhynchus grandis*  
3 Epialtidae: *Pelia tumida*  
3 Epialtidae: *Pugettia dalli*  
2 Epialtidae: *Pugettia producta*  
1 Epialtidae: *Scyra acutifrons*  
4 Epialtidae: *Taliepus nuttallii*  
1 Eulimidae: *Eulima raymondi*  
1 Facelinidae: *Hermisenda crassicornis*  
2 Fissurellidae: *Fissurella volcano*  
1 Flabelligeridae: *Piromis harrisae*  
1 Flabellinidae: *Flabellinopsis iodinea*  
1 Flabellinidae: *Flabellinopsis trilineata*  
1 Galeommatidae: *Chlamydoconcha orcutti*  
1 Gorgoniidae: *Leptogorgia chilensis*  
2 Grapsidae: *Pachygrapsus crassipes*  
2 Glyceridae: *Glycera americana*  
1 Glyceridae: *Glycera macrobranchiata*  
2 Hermaeidae: *Hermaea oliviae*  
2 Hesionidae: *Oxydromus pugettensis*  
4 Hippidae: *Emerita analoga*  
1 Hippolytidae: *Hippolyte californiensis*  
1 Hippolytidae: *Hippolyte clarki*  
2 Idoteidae: *Colidotea rostrata*  
1 Idoteidae: *Idotea urotoma*  
1 Idoteidae: *Pentidotea stenops*  
1 Idoteidae: *Pentidotea wosnesenskii*  
2 Idoteidae: *Synidotea harfordi*  
1 Inachoididae: *Pyromaia tuberculata*  
1 Ischnochitonidae: *Stenoplax conspicua*  
2 Ischyroceridae: *Ericthonius brasiliensis*  
3 Lepraliellidae: *Celleporaria*  
2 Lepraliellidae: *Celleporaria brunnea*  
2 Ligiidae: *Ligia occidentalis*  
2 Limidae: *Limaria hemphilli*  
1 Lineidae: *Lineus pictifrons*  
1 Lithodidae: *Glyptolithodes cristatipes*  
1 Lithodidae: *Lopholithodes foraminatus*  
1 Littorinidae: *Littorina plena*  
2 Lottidae: *Lottia gigantea*  
1 Lottidae: *Lottia scabra*  
1 Lottidae: *Lottia strigatella*  
1 Loveniidae: *Lovenia cordiformis*  
1 Luidiidae: *Luidia armata*  
2 Luidiidae: *Luidia foliolata*

- 1 Maeridae: Quadrimaera carla  
1 Membraniporidae: Membranipora membranacea  
1 Munididae: Munida hispida  
1 Muricidae: Roperia poulsoni  
1 Mytilidae: Leiosolenus plumula  
1 Mytilidae: Mytilisepta bifurcata  
1 Naticidae: Calinaticina oldroydii  
3 Nereididae: Gymnonereis crosslandi  
2 Nereididae: Platynereis bicanaliculata  
1 Octopodidae: Octopus rubescens  
2 Oenonidae: Arabella semimaculata  
2 Olivellidae: Olivella biplicata  
1 Onchidorididae: Acanthodoris brunnea  
1 Onchidorididae: Acanthodoris lutea  
1 Onchidorididae: Acanthodoris rhodoceras  
3 Ophiactidae: Ophiactis simplex  
1 Ophiodermatidae: Ophioderma panamensis  
2 Ophionereididae: Ophionereis eurybrachiplax  
1 Ophiopteridae: Ophiopteris papillosa  
1 Ophiotrichidae: Ophiothrix spiculata  
2 Ophiuridae: Ophiura luetkenii  
1 Opisthoteuthidae: Opisthoteuthis  
1 Orbiniidae: Leitoscoloplos pugettensis  
1 Orbiniidae: Naineris  
2 Orbiniidae: Phylo ornatus  
2 Orbiniidae: Scoloplos acmeceps  
1 Pacificincolidae: Primavelans insulta  
2 Paguridae: Phimochirus californiensis  
1 Palaemonidae: Palaemon macrodactylus  
2 Pandalidae: Pandanus platyceros  
2 Panopeidae: Lophopanopeus  
1 Parthenopidae: Latulambrus occidentalis  
1 Penaeidae: Farfantepenaeus californiensis  
1 Penaeidae: Parapenaeus longirostris  
1 Phidoloporidae: Phidolopora pacifica  
2 Philinidae: Philine auriformis  
2 Photidae: Gammaropsis thompsoni  
1 Phyllodocidae: Pterocirrus burtoni  
2 Pilumnidae: Pilumnus spinohirsutus  
1 Pleurobranchaeidae: Pleurobranchaea californica  
1 Podoceridae: Podocerus fulanus  
1 Poecilochaetidae: Poecilochaetus martini  
4 Pollicipedidae: Pollicipes polymerus  
3 Polyceridae: Limacia mcdonaldi  
2 Polyceridae: Polycera atra  
1 Polyceridae: Triopha maculata  
2 Polynoidae: Halosydna brevisetosa  
3 Porcellanidae: Petrolisthes cabrilloi  
1 Porcellionidae: Porcellionides pruinosus  
2 Portunidae: Portunus xantusii  
1 Potamididae: Cerithideopsis californica

- 1 Psolidae: *Lissothuria nutriens*  
1 Pycnogonidae: *Pycnogonum stearnsi*  
2 Sabellariidae: *Neosabellaria cementarium*  
1 Sabellariidae: *Phragmatopoma californica*  
1 Sabellidae: *Paradialychone ecaudata*  
1 Scalibregmatidae: *Scalibregma californicum*  
1 Scalpellidae: *Hamatoscalpellum californicum*  
1 Schizasteridae: *Brisaster latifrons*  
1 Sepiolidae: *Rossia pacifica*  
2 Sicyoniidae: *Sicyonia ingentis*  
1 Sicyoniidae: *Sicyonia penicillata*  
2 Sphaeromatidae: *Dynoides elegans*  
1 Sphaeromatidae: *Paracerceis sculpta*  
1 Spionidae: *Laonice cirrata*  
1 Spionidae: *Parapronospio alata*  
1 Stenothoidae: *Stenothoe estacula*  
2 Stichopodidae: *Apostichopus californicus*  
3 Strongylocentrotidae: *Mesocentrotus franciscanus*  
3 Strongylocentrotidae: *Strongylocentrotus fragilis*  
2 Styelidae: *Botrylloides violaceus*  
1 Styelidae: *Polyandrocarpa*  
1 Styellidae: *Styela montereyensis*  
2 Styellidae: *Styela plicata*  
1 Syllidae: *Amblyosyllis speciosa*  
1 Syllidae: *Epigamia noroi*  
2 Syllidae: *Megasyllis nipponica*  
1 Syllidae: *Myrianida pentadentata*  
1 Tanaididae: *Zeuxo normani*  
1 Tegulidae: *Norrisia norrisii*  
1 Tegulidae: *Tegula aureotincta*  
3 Tegulidae: *Tegula eiseni*  
1 Terebellidae: *Pista brevibranchiata*  
1 Tetractitidae: *Tetraclita rubescens*  
1 Toxopneustidae: *Lytechinus pictus*  
1 Tubulanidae: *Tubulanus sexlineatus*  
1 Turbinidae: *Megastraea undosa*  
3 Upogebidae: *Upogebia macginitieorum*  
2 Valenciniidae: *Baseodiscus*  
2 Velutinidae: *Hainotis sharonae*  
1 Veneridae: *Venerupis philippinarum*  
2 Vermetidae: *Thylacodes squamigerus*  
2 Virgulariidae: *Stylatula elongata*  
2 Watersiporidae: *Watersipora subtorquata*  
2 Xanthidae: *Paraxanthias taylori*  
1 Zoroasteridae: *Myxoderma platyacanthum*

---

## **APPENDIX II: FULL METHODS**

### **Contamination precautions**

False positives may result from contamination during eDNA collection as well as during DNA extraction and amplification. To prevent contamination in both field sampling and lab work, we followed precautions outlined in Goldberg et al. (2016). Specifically, personnel wore nitrile gloves at all times during field and lab sample processing. Prior to eDNA sample collection, we sterilized all containers, supplies, and work surfaces with 10–30% bleach solution followed by 30 min ultraviolet light (UV) treatment. We conducted all eDNA filter extractions and PCR preparations in an AirClean 600 PCR Workstation (Creedmoor, NC, USA) located in a clean room dedicated to DNA extractions and PCR preparations at the University of California, Los Angeles (UCLA). Before and after use, we cleaned the AirClean 600 PCR Workstation and pipettes with 10–30% bleach followed by a 30 min UV treatment. We used filtered pipette tips for all pre- and post-PCR procedures; these and other consumables were decontaminated by a 30 min UV treatment. In addition to the above, we included negative controls at all sample processing steps including field blanks, extraction blanks, and PCR blanks, we also used two positive controls to test for index hopping.

### **Environmental DNA extractions, amplification and sequencing**

At UCLA, all eDNA filter extractions and PCR preparations were conducted in an AirClean 600 PCR Workstation (Creedmoor, NC, USA) in a clean room dedicated to DNA extractions. The AirClean 600 PCR Workstation and pipettes were decontaminated before and after use with 10% bleach followed by a 20 min ultraviolet light (UV) treatment. Filtered pipette tips were used for all protocols. There were negative controls (extraction blanks and PCR blanks) for all steps in the sample processing.

Sterivex filters were extracted with a DNAeasy Tissue and Blood Qiagen Kit (Spens et al. 2017) protocol. Proteinase K and ATL buffer was added to the filter cartridge and incubated overnight at 56° C. DNA was amplified with primer sets with Illumina Nextera adapter modifications for use with Nextera XT (indexes (Product Number FC-131-2001 through -2004; 5200 Illumina Way, San Diego, CA 92122). The four primer sets used were: (1) MiFish Universal Teleost 12S primer (176bp) (Miya et al. 2015); (2) MiFish Universal Elasmobranch 12S primer (186bp) (Miya et al. 2015); (3) metazoan COI primers (313bp) (Leray et al. 2013); and (4) metazoan 16S primers (~115bp) (Kelly et al. 2016). PCR amplifications in triplicate for each primer pair contained 12.5 µL QIAGEN Multiplex Taq PCR 2x Master Mix, 6.5 µL dH<sub>2</sub>O, 2.5 µL of each primer (2 µmol/L), and 1 µL DNA template (25 µL total reaction volume). PCR thermocycling employed a touchdown profile with an initial denaturation at 95° C for 15 min to activate the DNA polymerase followed by 13 cycles with a denaturation step at 94° C for 30 sec, an annealing step with temperature starting at 69.5° C for 30 sec (temperature was decreased by 1.5° C every cycle until 50° C was reached), and an extension step at 72° C for 1 min. An additional 35 cycles were then carried out at an annealing temperature of 50° C using the same denaturation and extension steps above, followed by a final extension at 72° C for 10 min. All PCRs

included a negative control, where molecular grade water replaced the DNA template. All PCR products were run on 2% agarose gels to ensure amplification success and correct product size.

To prepare PCR products for sequencing, we pooled triplicate PCR reactions using 5 µL volume from each PCR. Pooled samples were cleaned using Serapure magnetic beads (Faircloth and Glenn 2014) and cleaned PCR product concentrations were quantified using the high sensitivity Quant-iT™ dsDNA Assay Kit (Thermofisher Scientific, Waltham, MA, USA) on a Victor3™ plate reader (Perkin Elmer Waltham, MA, USA). Sample DNA libraries were prepared using the Nextera Index Kits (Illumina, San Diego, CA, USA) and KAPA HiFi HotStart Ready Mix (Kapa Biosystems, Wilmington, MA, USA). This second indexing PCR was performed using a 25 µL reaction mixture containing 12.5 µL of Kapa HiFi Hotstart Ready mix, 0.625 µL of primer i7, 0.625 µL of primer i5, and 10 ng of template DNA, and used the following thermocycling parameters: denaturation at 95° C for 5 min, 5 cycles of denaturation at 98° C for 20 sec, annealing at 56° C for 30 sec, extension at 72° C for 3 min, followed by a final extension at 72° C for 5 min. All indexed PCR products were run on 2% agarose gels to ensure successful PCR and correct product size. Resulting libraries were bead cleaned and quantified as described above. Finally, we pooled indexed libraries by barcode in equimolar concentration, and then sequenced each of the four libraries on a MiSeq at the Technology Center for Genomics & Bioinformatics (University of California Los Angeles, CA, USA), using Reagent Kit V3 with 20% PhiX added to all sequencing runs.

### Bioinformatic analysis and taxonomic determination

We used the Anacapa Toolkit (Curd et al. 2019) for amplicon sequence variant parsing, taxonomic assignment, and quality control. Taxonomic assignments of eDNA metabarcoding reads require a comprehensive sequence reference library. To generate these libraries for fish and invertebrates, we queried NCBI GenBank 28 May 2019 for all available fish and invertebrate sequences that overlapped our primer regions. Three libraries were constructed: (1) 12S for fish, rays, and sharks; (2) COI; and (3) 16S for metazoans including invertebrates. We processed sequences using the default parameters and then assigned taxonomy using reference libraries generated using CRUX (Constructing Reference libraries Using eXisting tools) (Curd et al. 2019) These libraries are limited to the taxa they contain. Taxa for which these barcode regions are not available cannot be matched with eDNA metabarcoding data.

The quality control step of the Anacapa Toolkit trims extraneous adapter sequences used to identify each unique sample, removes low quality reads, and sorts reads by metabarcode primer sequence. A key advantage of the Anacapa Toolkit is that it can simultaneously processes raw fastq reads for samples with single or multiple metabarcode targets generated on Illumina MiSeq machines. It is also not required that all samples contain reads for each metabarcode, thus allowing users to combine multiple projects or targets on the same sequencing run while only running the pipeline once.

The amplicon sequence variant (ASV) parsing step uses DADA2 (Callahan et al. 2016) to derePLICATE our metabarcodes. ASVs are a novel solution to identifying biologically informative unique

sequences in metabarcoding samples that replaces the operational taxonomic unit (OTU) framework. Unlike OTUs, which cluster sequences using an arbitrary sequence similarity (e.g. 97%, which corresponds to a gross average genetic sequence difference across all of life), ASVs are unique sequence reads determined using Bayesian probabilities of known sequencing error. These unique sequences can differ by as little as two base pairs, providing improved taxonomic resolution and an increase in observed diversity (Callahan et al. 2016, Amir et al. 2017).

Next the Anacapa toolkit module assigns taxonomy to ASVs using Bowtie 2 (Langmead and Salzberg 2012) and a Bowtie 2-specific Bayesian Least Common Ancestor (BLCA) algorithm. All ASVs are first globally aligned against the CRUX database using Bowtie 2. Any ASV that fails to align is then aligned locally. The best hits (the top 100 Bowtie 2 returns) are then processed with the BLCA script to assign taxonomy. The Bowtie 2 BLCA algorithm was adapted from <https://github.com/qun-fengdong/BLCA>. BLCA uses pairwise sequence alignment to calculate sequence similarity between query sequences (a given ASV sequence found in the environment) and reference library hits. Taxonomy is assigned based on the lowest common ancestor of multiple reference library hits for each query sequence. The reliability of each taxonomic assignment is then evaluated through bootstrap confidence scores (Gao et al. 2017). Scores are based on Bayesian posterior probability which quantify the similarity of reference barcode sequences to the query sequence. The higher the similarity between the database barcode sequence and the query sequence, the greater the contribution to the taxonomic assignment of the query. Ultimately, this method provides a strong probabilistic basis for evaluating taxonomic assignments with bootstrap confidence scores.

## Bioinformatic Decontamination

After processing the raw sequence reads through the Anacapa Toolkit, the resulting species community tables were transferred into R (R Core Team 2016) for subsequent decontamination and downstream data analysis. The raw species community table needs to be decontaminated to eliminate potential field contamination, lab contamination, and sequence index hopping (Goldberg et al. 2016, Costello et al. 2018). Field and lab contamination can arise because of inadequate sterile procedures, careless laboratory work, and reagents, particularly enzymes which are generated from living organisms (Goldberg et al. 2016). Sequence index hopping occurs when the DNA index tag used to label each unique sample chemically swaps with the DNA index tag of another sample, leading to cross contamination of species between samples. However, to address these sources of contamination, we followed rigorous sterile procedures and sequenced our data on platforms that limit index hopping.

In addition to these precautionary steps, we also implemented a decontamination procedure that eliminates any remaining sources of contamination (Kelly et al. 2018 [supplemental methods], 2019 [supplemental methods], Gallego 2019). We implemented two steps of decontamination: the first identifying and removing sequences arising from index hopping and second identifying and removing sequences arising from negative controls. Briefly, we accomplish the first step by estimating the rate of index hopping of each ASV into negative controls and use this as a baseline for removing likely molecular recombination of index sequences. Second, we identify the likely origin

of each ASV from either samples or control by comparing the total number of reads, prevalence (number of samples occurred), the proportion of reads, and proportion of prevalence of each ASV in both samples and controls. We then compare whether each of these metrics was higher for a given ASV in samples or controls. We eliminate any ASVs with metrics greater in controls than samples. We note this metric is a highly conservative approach to decontamination, but ensures high confidence interpretation of results

### **Normalizing eDNA reads across samples**

Relating the number of sequence reads in a sample to the number of individuals of a taxon is extremely challenging. Not only do different organisms shed DNA at different rates, but the process of PCR amplification is known to have sequence-specific biases. However, it is possible to address some of those known biases in ways that make it plausible to explore eDNA's ability to yield insights into taxon populations (Taberlet et al. 2018).

We transformed our data into an “eDNA index” (Kelly et al. 2019). This metric assumes that PCR biases originate from template-primer interactions which remain constant across eDNA samples. Intuitively, therefore, it should be possible to infer information about taxon abundance by using appropriate normalizations of the counts of reads in each sample. The eDNA index transformation is conducted by first normalizing all reads for a particular sequence by the total number of reads in each sample, then scaling those proportions to the largest observed proportion for that sequence across all samples. This results in a sequence-specific (species-specific) scaling between 0 to 1, where 1 is the sample with the highest number of reads for a given species and 0 is the least.

### **Barcode reference sequences generated for this project**

As described above, taxa for which there is no reference sequence cannot be identified using eDNA. Therefore, to augment publicly-available bony and cartilaginous fish reference sequences, an additional 16 fish barcode references were generated for this project, based on fish species lists generated by Biosurveys.

For invertebrates, public databases are far less complete than they are for vertebrates. Hence for this project, we committed to sequencing an additional 200 reference barcode sequences for local invertebrates. Voucher specimen sampling (Department of Fish and Wildlife Permit CA-004668) was not limited to benthic trawls, but also included other sampling methods (e.g. beach seines and bottom grabs) from Los Angeles County localities. Specimens were identified to the level of species, DNA extracted, amplified, sequenced, and added to the previously generated publicly available pubbarcode dataset. Vouchered and databased specimens are permanently housed in the NHMLA collections.

Genomic DNA from fish and invertebrates was isolated using either: (1) the DNeasy® Blood and Tissue Kit (Qiagen, Valencia, CA, USA) following the manufacturer’s protocol or (2) using an AutoGenprep 965 using the AutoGen standard Mouse Tail (Animal Tissue) protocol (Holliston, MA, USA). For fishes, MiFish 12S primers were used (Miya et al. 2015), and for invertebrates, metazoan

COI (Geller et al. 2013) primers were used. PCR products were purified using ExoSAP-IT (Affymetrix, Cleveland, OH, USA) and sequenced in both primer directions. Forward and reverse sequences were trimmed at the ends and assembled into contigs in Geneious version 8.1.9 (Kearse et al. 2012) and aligned using MUSCLE (Edgar 2004).

## R analysis methods — Statistical methods

In order to compare species identified with eDNA metabarcoding and visual survey methods, we compared taxonomic rank (phylum, class, order, family, genus, and species) of organisms identified from Biosurvey methods, trawls, and eDNA metabarcoding in R (R Core Team 2016). We specifically investigated the overlap and non-overlap of species detected by eDNA metabarcoding and visual survey methods. In addition, we compared the overlap and separation between species identified in Biosurvey and trawls against eDNA reference databases to investigate possible explanations for why eDNA metabarcoding approaches failed to identify species detected in Biosurvey data. In addition, we compared the number of invasive species in the CAL-NEMO database (Fofonoff et al. 2019) detected across all Biosurveys and eDNA metabarcoding efforts.

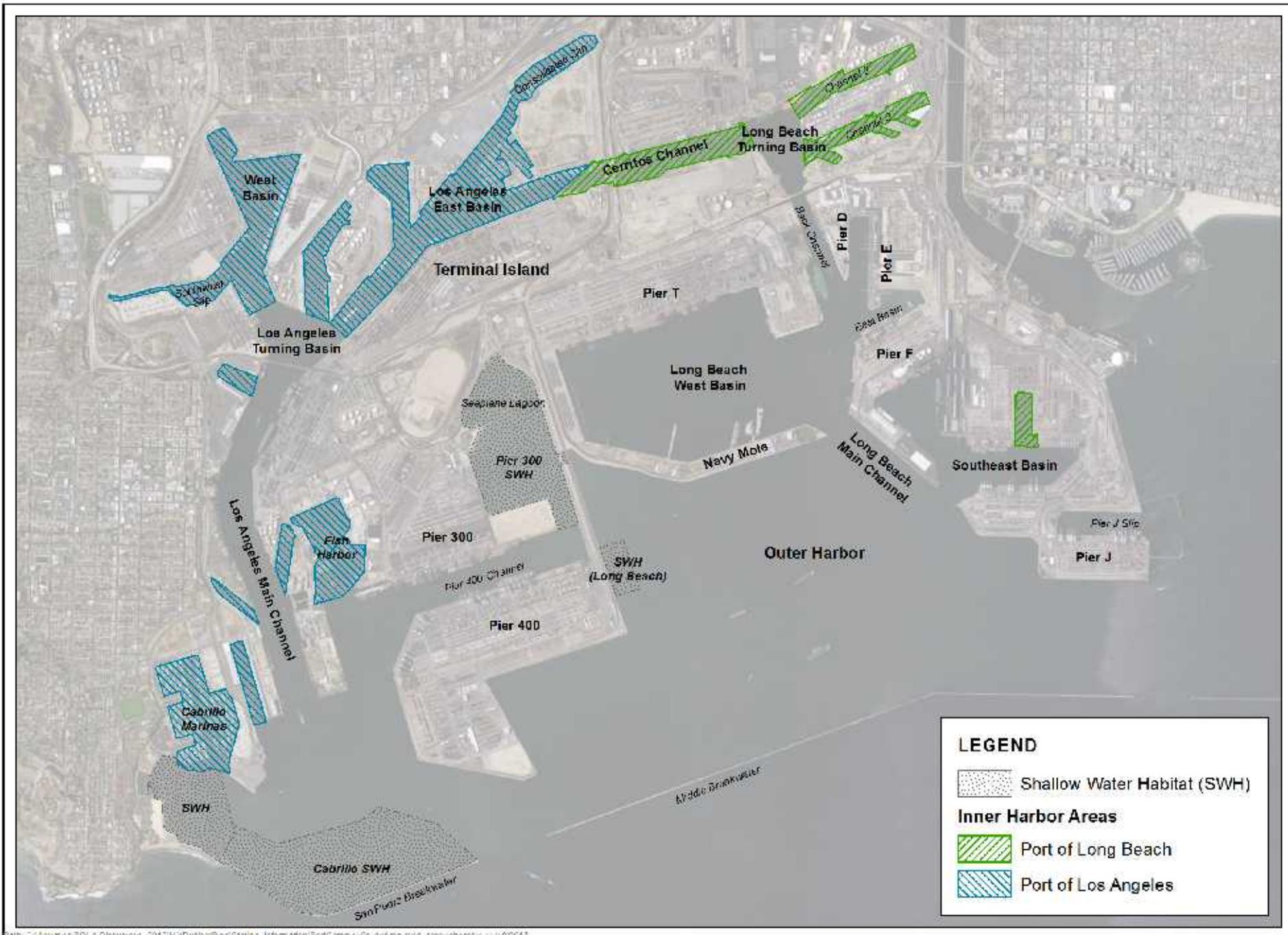
In order to investigate the relationship between eDNA index scores and trawl biomass and relative abundance, we fit linear regressions between both metrics in R. We calculated and reported r-squared and equation values for each fit.

We used the *phyloseq* (McMurdie and Holmes 2013) and *ranacapa* (Kandlikar et al. 2018) R packages to import and manipulate the eDNA metabarcoding data in R. We compared the total number of species found at each site from eDNA metabarcoding analysis using Analysis of Variance (ANOVA) and Tukey post-hoc tests using the R package *vegan* (Oksanen et al. 2019). We also calculated species accumulation curves to conduct a power analysis of the number of additional species found for each additional replicate across each site using the *iNext* (Chao et al. 2014, Hsieh et al. 2019) and *vegan* R packages.

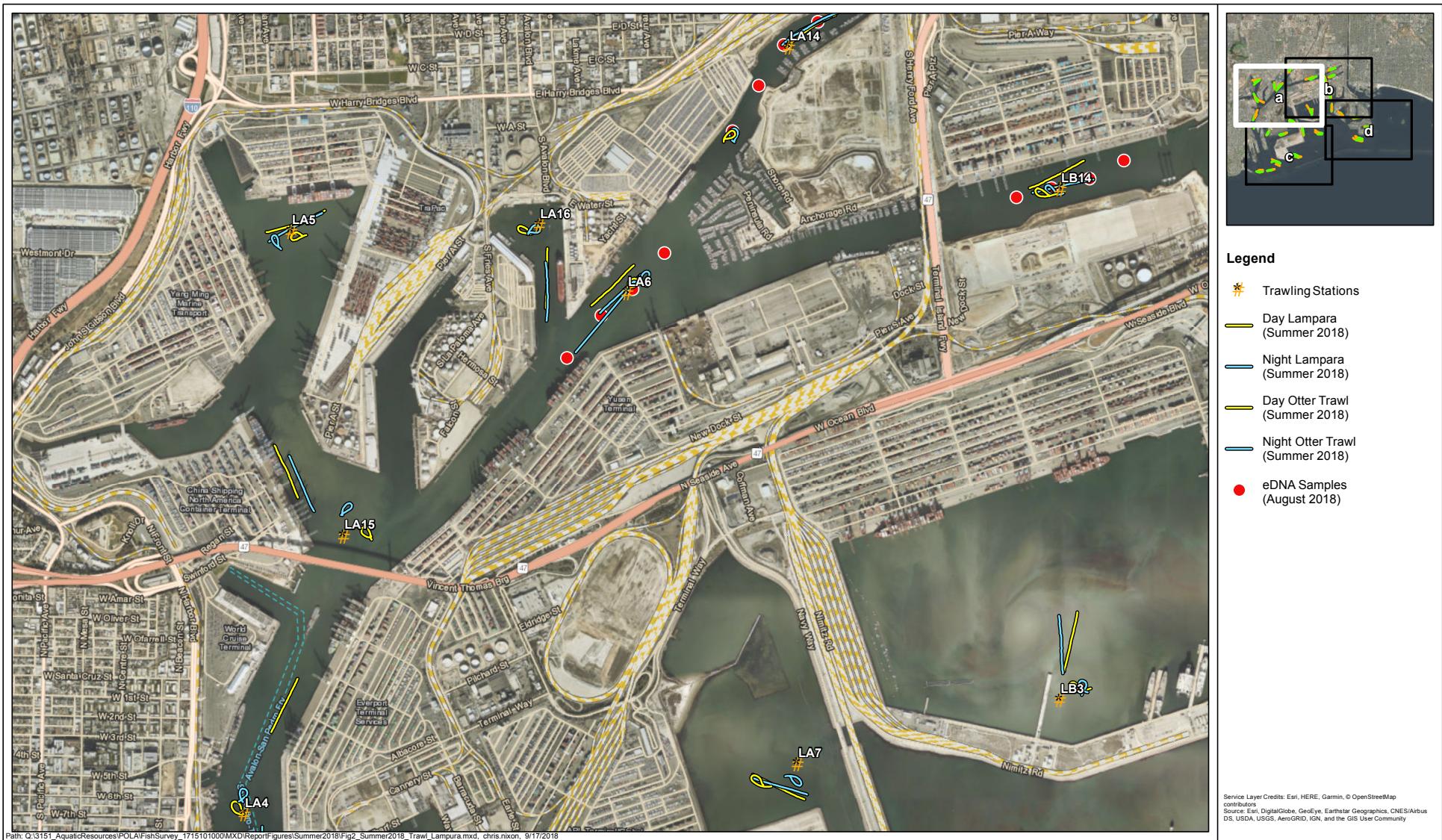
We calculated the pairwise similarity of all 87 eDNA samples using the Bray Curtis similarity index through the *vegan* package. We then apportioned variation between the pairwise similarity of each sample against site, station, data collected, and volume filtered using a PERMANOVA and Betadisp analysis through *vegan*. We visualized the apportioned variance using the *treemapify* R package (Wilkins 2019) and a NMDS ordination through *phyloseq* and *vegan* packages.

---

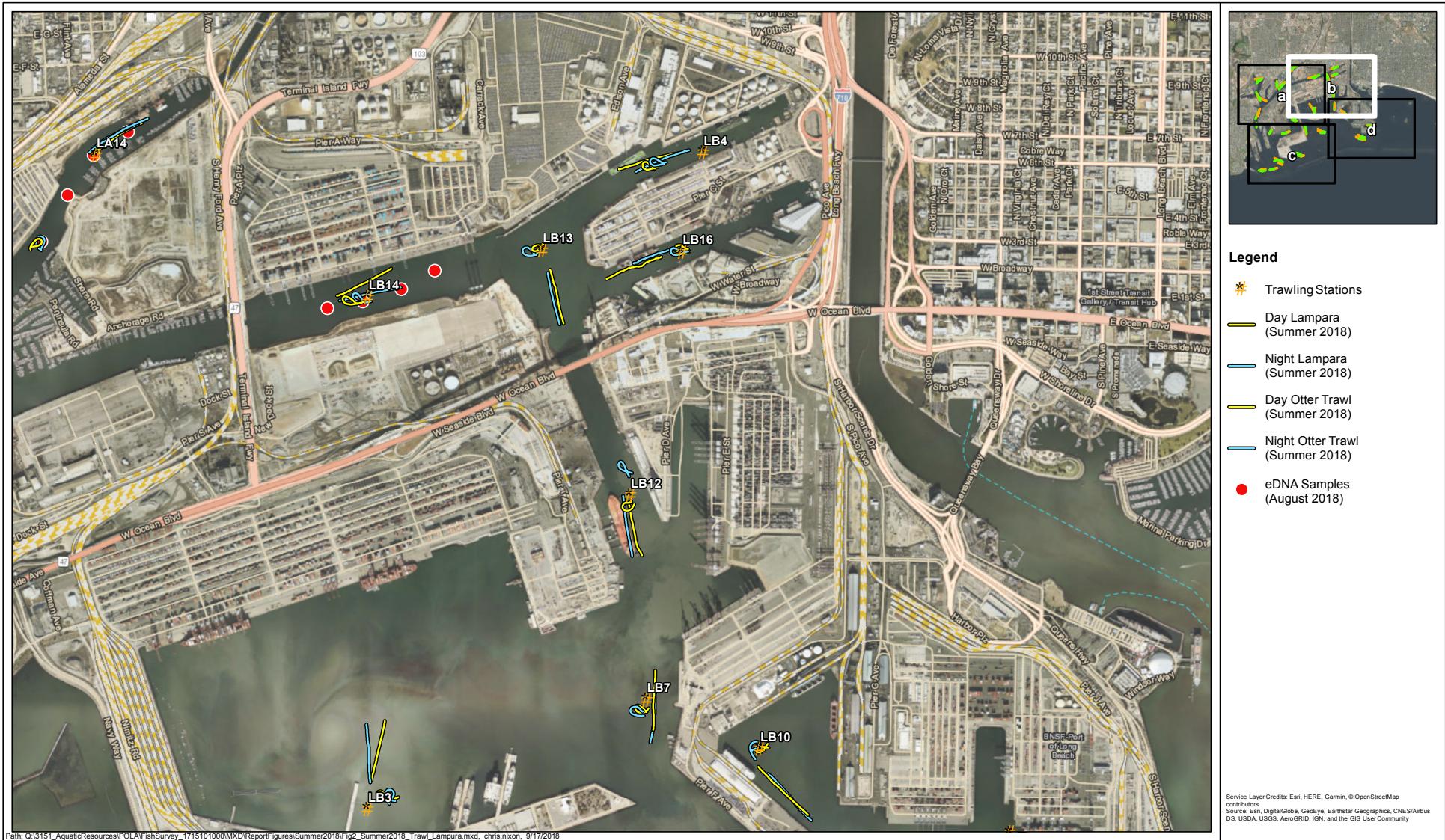
### **APPENDIX III: MAPS OF SAMPLING STATIONS AND LOCATIONS**



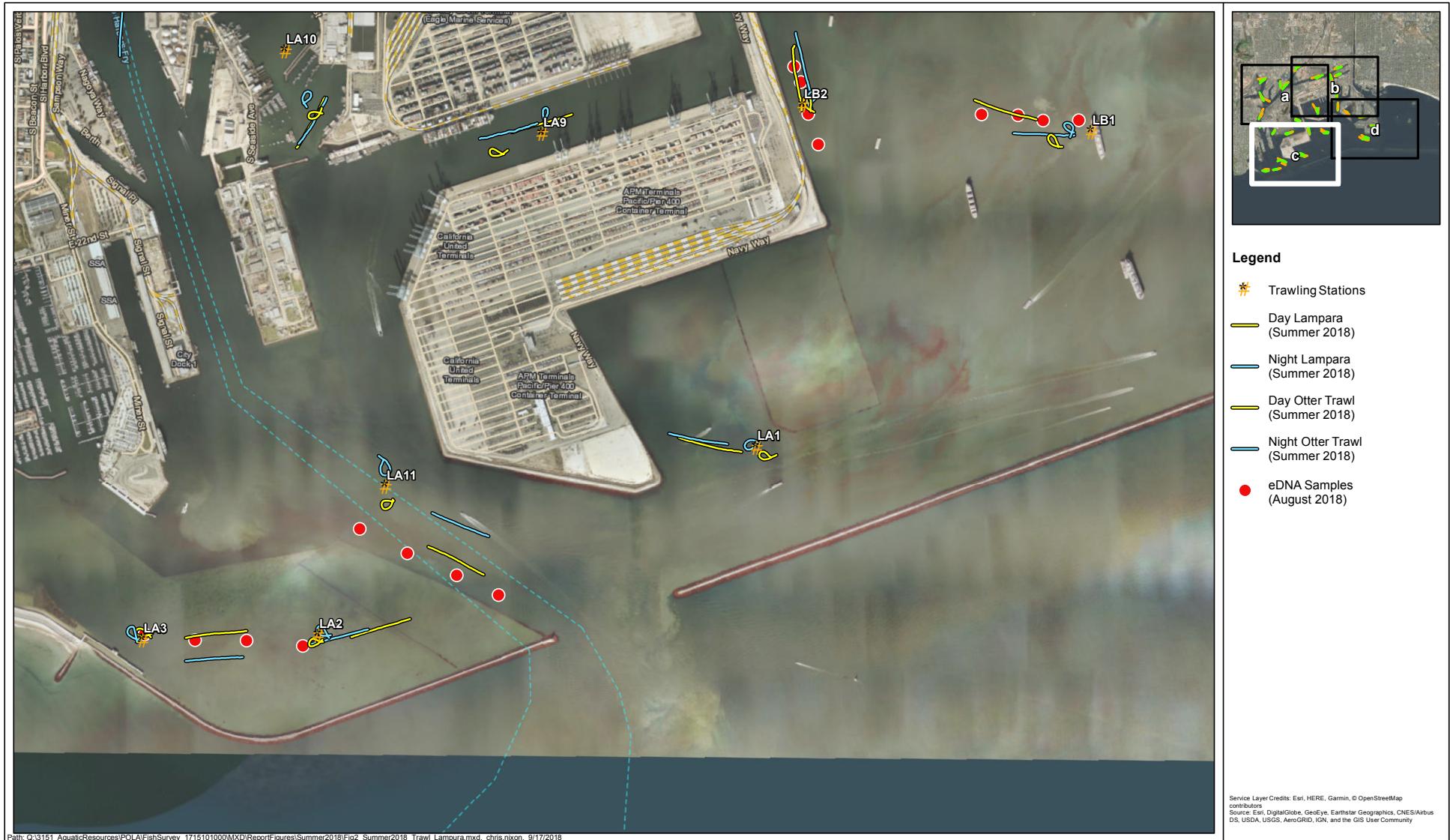
The Port Complex Study Area  
Ports of Los Angeles and Long Beach  
San Pedro Bay, CA



Path: Q:\3151\_AquaticResources\POLA\fishSurvey\_171510\1000\MDX\Report\figures\Summer2018\Fig2\_Summer2018\_Trawl\_Lampara.mxd, chris.nixon, 9/17/2018



Summer 2018 Otter Trawl and Lampara Net Deployments  
2018 Biological Surveys of the Port of Los Angeles and Long Beach  
San Pedro Bay, CA

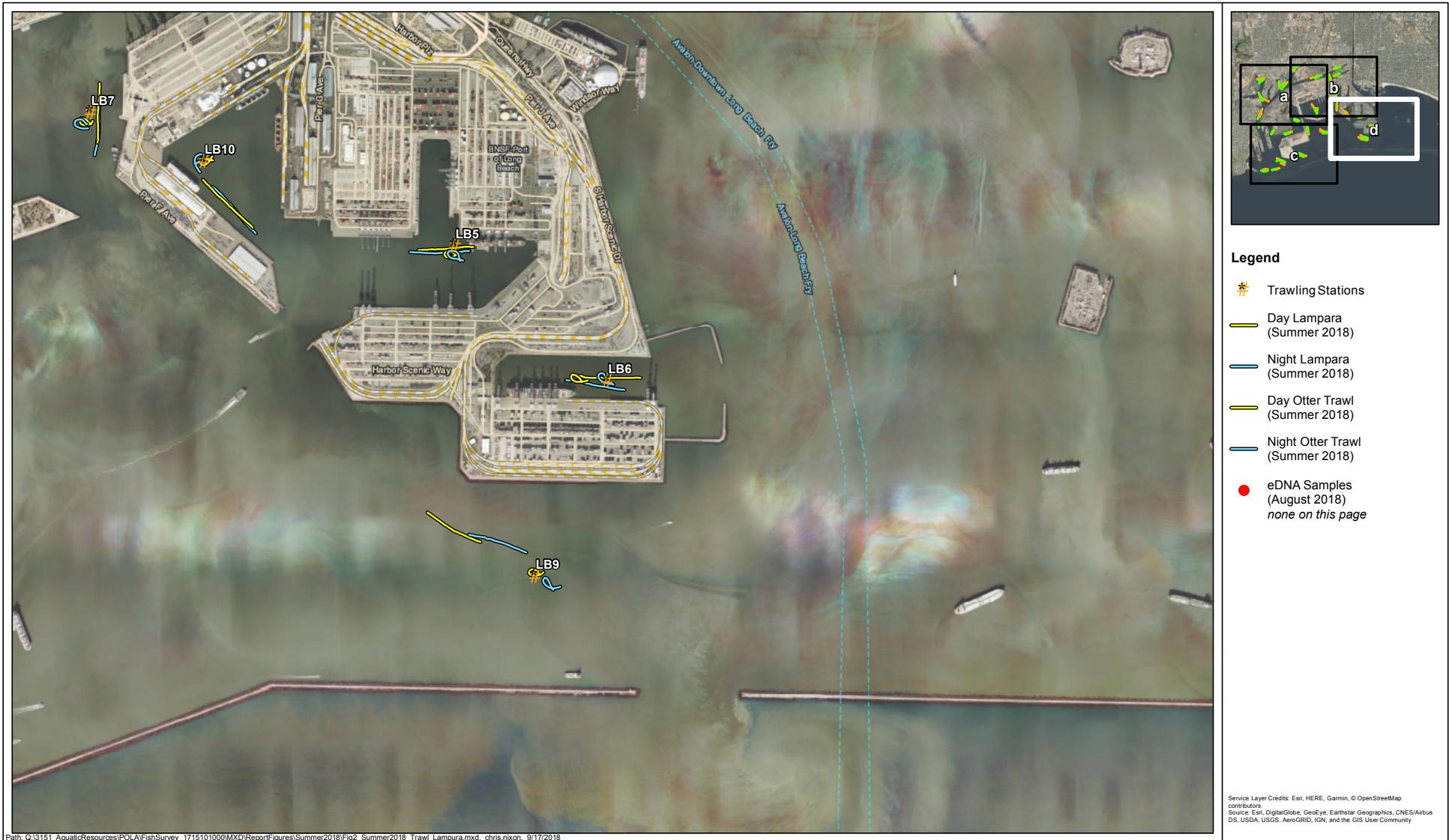


Path: Q:\3151\_AquaticResources\POLA\FishSurvey\_1715101000\MD\Report\Figures\Summer2018\Fig2\_Summer2018\_Trawl\_Lampara.mxd, chris.nixon, 9/17/2018

Summer 2018 Otter Trawl and Lampara Net Deployments  
2018 Biological Surveys of the Port of Los Angeles and Long Beach  
San Pedro Bay, CA

amec foster wheeler

0 1,500 Feet  
1 inch = 1,500 feet



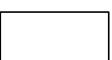
Path: Q:\3151\_AquaticResources\POLAFishSurvey\_1715101000\MD\Report\ReportFigures\Summer2018\Fig2\_Summer2018\_Trawl\_Lampara.mxd, chris.nixon, 9/17/2018

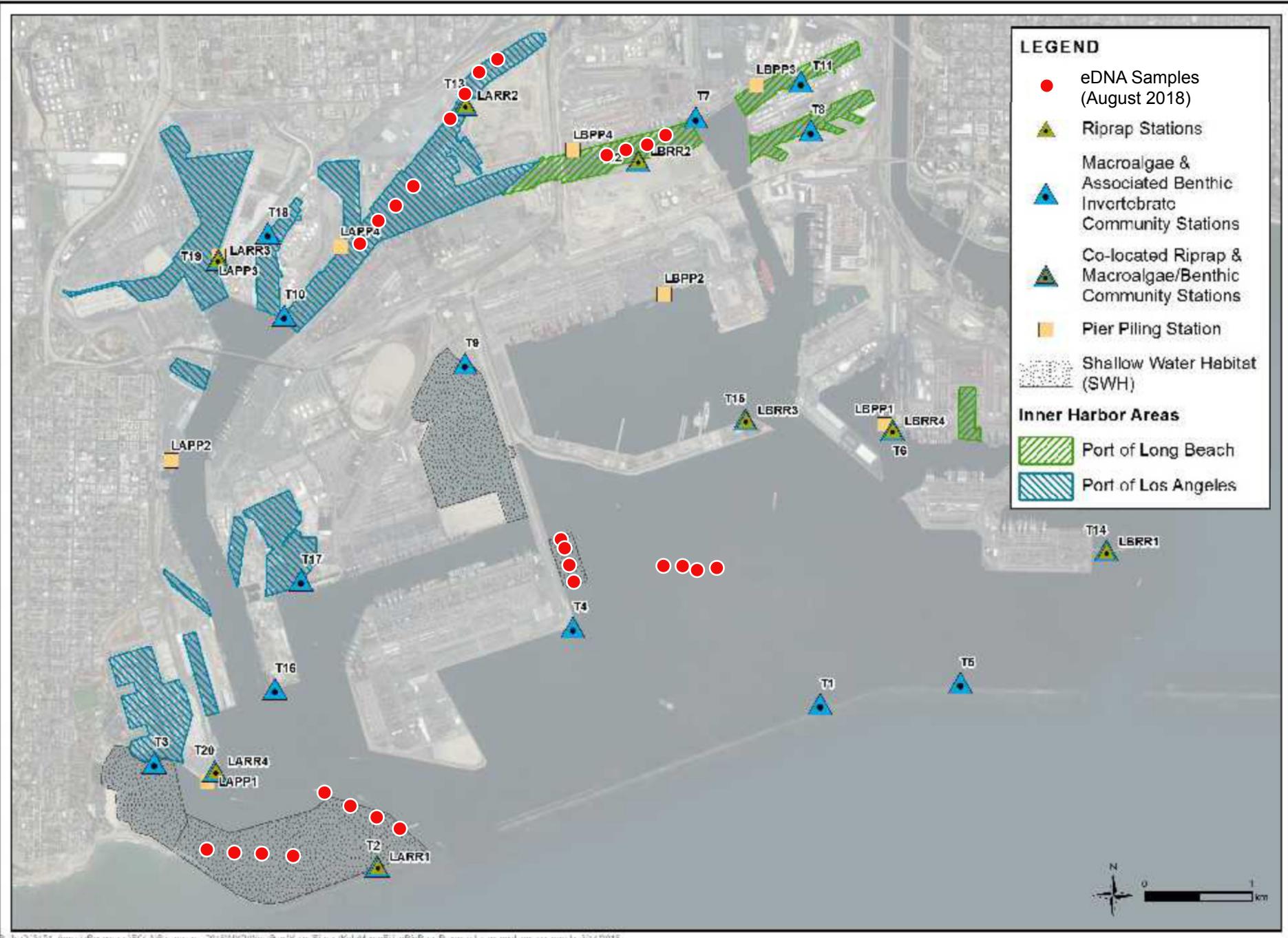
**Summer 2018 Otter Trawl and Lampara Net Deployments**  
**2018 Biological Surveys of the Port of Los Angeles and Long Beach**  
**San Pedro Bay, CA**

amec foster wheeler



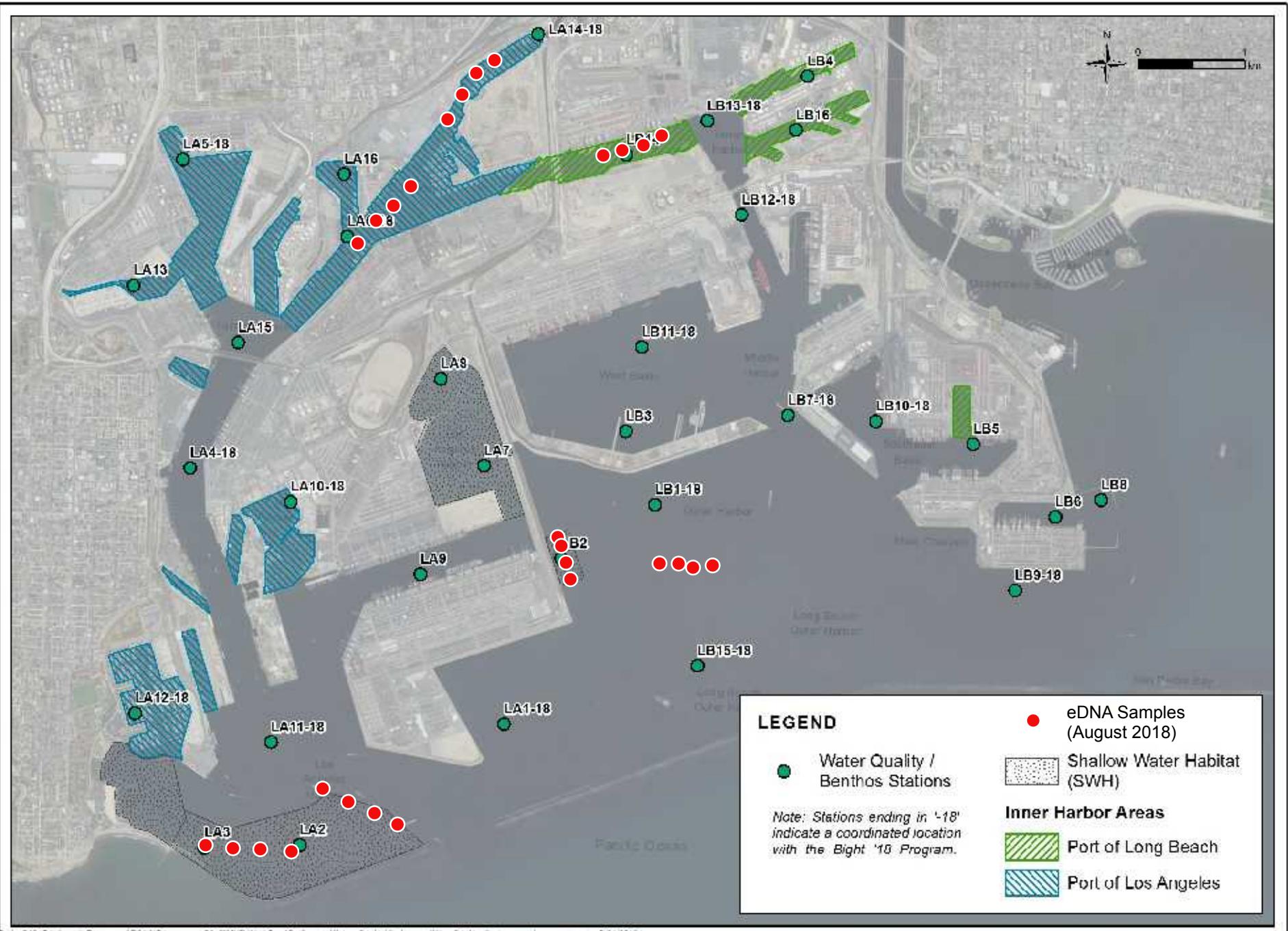
1 inch = 1,500 feet  
 0 1,500 Feet





B:\h\03157\AquaticResource\2018\PC\A\8.Csv, v. 201804031449 - Port of Los Angeles/Marina/Riprap/Pier Piling, eDNA, SWH.mxd, created on 2014/05/05

**Riprap, Pier Piling, Macroalgae, and Associated Benthic Invert. Community Stations  
2018 Biological Surveys of the Ports of Los Angeles and Long Beach  
San Pedro Bay, CA**



B.1.02157\_AquaticResource/POCALB/2018/04/04/PortofLongBeach\_Sediment\_WaterQuality\_Euro\_4\_7.2\_v01a.vrt 2/14/2018

**Sediment and Water Quality Stations**  
**2018 Biological Surveys of the Ports of Los Angeles and Long Beach / Bight '18**  
**San Pedro Bay, CA**