Message in a Bottle: Archived DNA Reveals Marine Heatwave-Associated Shifts in Fish Assemblages

Zachary Gold1,2\*, Ryan P. Kelly3, Andrew Olaf Shelton2, Andrew R. Thompson4, Kelly D. Goodwin4,5, Ramón Gallego2, Kim M. Parsons2, Luke R. Thompson5,6, Dovi Kacev7, Paul H. Barber8

1 Cooperative Institute for Climate, Ocean, & Ecosystem Studies, UW, Seattle, WA

2 Northwest Fisheries Science Center, NMFS/NOAA, Seattle, WA

3 School of Marine and Environmental Affairs, UW, Seattle, WA

4 Southwest Fisheries Science Center, NMFS/NOAA, La Jolla, CA

5 Ocean Chemistry and Ecosystems Division, Atlantic Oceanographic and Meteorological Laboratory, Miami, FL

6 Northern Gulf Institute, Mississippi State University, Mississippi State, MS

7 Scripps Institution of Oceanography, UCSD, La Jolla

8 Department of Ecology and Evolutionary Biology, UCLA, Los Angeles, CA

\*Corresponding author Email: [zachary.gold@noaa.gov](mailto:zachary.gold@noaa.gov)

**Introduction**

The CalCOFI program (<https://calcofi.com/>) serves to provide fisheries-independent ecosystem assessments of fish assemblages in the Southern California Current and has provided decades of data on ichthyoplankton assemblages (Gallo et al., 2019). The current CalCOFI surveys sample four times per year from the U.S. Mexican Border to Monterey Bay (Gallo et al., 2019). We used this rich sample archive to interrogate ichthyoplankton assemblages from 1996-2019. This supplemental material provides additional details on the methods to support the main findings and conclusion of the manuscript.

**Methods**

*Study Design*

To evaluate the efficacy of metabarcoding methods used to analyze ethanol preserved samples and investigate potential changes in the ichthyoplankton assemblages over decadal scales, we identified ichthyoplankton by metabarcoding and microscopy in ethanol-preserved samples collected over two decades (1996,1998-2019; Figure 1) of spring CalCOFI cruises (5, 27, 35, 38). We note that samples collected in 1997 were stored in <50% ethanol and were discarded due to failed preservation.

Samples were collected in late March or early April of each year (calcofi\_metadata\_analysis\_20210907.csv). Here we focus on spring samples because the majority of species in the California Current spawn in spring and historically the annual California Current Ecosystem Report has relied on the spring data (McClatchie et al., 2018; Thompson et al., 2022). This decision is supported by recent work using ichthyoplankton data across the full set of yearly CalCOFI cruises which found little evidence for phenological trends (McClatchie et al., 2018; Thompson et al., 2022), thus aiding the ability to look at impacts from the marine heatwave.

Archived spring ichthyoplankton samples were collected across four biogeographically dissimilar stations (up to 370 km apart) with variable water properties (McClatchie et al., 2016) over 2 decades (1996,1998-2019) (Nielsen et al., 2021; Thompson et al., 2019). The northernmost station was located offshore of Point Conception, CA within the California Current (34.14833˚N -121.1567˚W). The second station was located off San Nicholas Island, CA (33.32333 ˚N, -119.6667˚W) which experiences high variation in annual temperature depending on the respective strengths of the California Current and Southern California Counter Current. The third station was a southern coastal inshore station off San Diego, CA (32.84667˚N, -117.5383˚W) characterized by relatively warmer waters from the California Counter Current with seasonal (spring) upwelling of cool, nutrient-rich water. The fourth station was a southern offshore station (31.85000˚N, -119.5683˚W) characterized by sub-tropical oceanic waters (Figure 1).

At each station, oblique bongo net tows were conducted from 210 m to the surface using standard CalCOFI methods (Kramer et al., 1972; McClatchie, 2014; Thompson et al., 2012, 2017). Each side of the bongo net had a 0.71 m-diameter mouth opening and a net size of 0.505 mm mesh. Cod end contents of both bongo nets were preserved at sea. The starboard side was preserved in sodium borate-buffered 2% formaldehyde and the port side was preserved in Tris-buffered 95% ethanol. Ethanol was replaced after 24 hours to account for dilution from tissue water loss. Microscopy was conducted to identify species abundance from formaldehyde-preserved samples following standardized CalCOFI techniques (McClatchie, 2014) while metabarcoding was conducted on the ethanol in which port side samples were stored; consequently, we expected the contents of the paired samples to differ slightly as a function of sampling stochasticity.

*Metabarcoding Collection Isolation, Amplification, and Sequencing*

Prior to filtration, the ethanol-preserved samples were inverted three times and let rest for 30 minutes to resuspend and homogenize samples in the preservative. Filtration of ethanol from the port-side bongo samples was conducted in a pre-PCR clean room at the NOAA Southwest Fisheries Science Center within a biological safety cabinet in July 2019. The pre-PCR room had no previous post-PCR work conducted within and all surfaces and equipment were sterilized frequently with 10% bleach and 70% ethanol. The pre-PCR clean room was at ambient pressure and reasonable precautions to limit contamination were conducted including only wearing clean clothes that have not been exposed to labs with PCR product, no food brought into the lab, and gloves were exchanged regularly.

Ethanol preservative was filtered using a vacuum filtration manifold with Nalgene Analytical Test Filter Funnels (Thermofisher Scientific, Waltham, MA, USA) with the manufacturer’s 0.45 µm filters replaced with 0.2 µm Durapore PVDF filters (Sigma Aldrich, St. Louis, MO, USA) using sterile forceps. Up to 125 mL of ethanol was then transferred from the preserved jars into the filter funnels using a 10 mL pipette, carefully avoiding any sample contents and thus preserving CalCOFI specimens for future research and analysis. Sample jars were refilled using freshly prepared tris-buffered ethanol before being returned to the collection archive. We included two negative controls to test for lab contamination by filtering 125 mL of molecular grade water. Filters were stored at -20˚C before DNA extraction.

Filters were extracted using the standard Qiagen DNAeasy Kit (Qiagen Inc., Valencia, CA, USA) in a pre-PCR molecular lab. Extracted DNA was amplified using the MiFish Universal Teleost primer sets to capture fish diversity (67).

Here, we highlight our decision to utilize the MiFish Universal Teleost *12S* primers. First, these primers (Table S1) have been rigorously validated for fish barcoding and shown to provide accurate taxonomic assignments for a broad range of fishes (Collins et al., 2021; Curd et al., 2019; Gold et al., 2021; Miya et al., 2015, 2020; Polanco F et al., 2021; Valsecchi et al., 2020). We recognize that there are limitations for this, and indeed all, metabarcoding primer sets (Deiner et al., 2017) which are forced to balance specificity [how well target species can be taxonomically resolved] against breadth [range of species across the tree of life that can be amplified] (Taberlet et al., 2018). Even a “gold standard” like the *16S* rRNA gene marker for prokaryotic sequences struggles with taxonomic assignment accuracy (Edgar, 2018), especially with short-read sequences. Although taxonomic resolution limitations and compromises remain for the *12S* target (Gold et al., 2021; Min et al., 2021), the taxonomic resolution has been improved and best practices for taxonomic classification have been identified through the development of a nearly comprehensive California Current Large Marine Ecosystem *12S* reference database along with a full factorial cross-validation analysis of bioinformatic approaches (Gold et al., 2021).

Second, there are no widely used or benchmarked CO1 metabarcoding primer sets for fish applications although CO1 barcoding is a common barcoding target. This is because a) the conserved nature of the locus across the tree of life which results in amplification of a broad array of taxa (Hastings & Burton, 2008; Leray et al., 2013), and b) the mismatch in high throughput sequencing platform length (max is paired-end 300 bp) and rate of CO1 evolution/accumulation of sequence differences between species (Deagle et al., 2014; Polanco F et al., 2021). In fact, these shortcomings were the original motivation for researchers to develop alternative fish metabarcoding loci targeting *12S* loci for fishes (23). Together, the research community has largely converged on the MiFish Universal Teleost *12S* primer set as standard practice for fish metabarcoding given its balance of high specificity and breadth (Miya et al., 2020). Thus we feel confident that the MiFish Universal Teleost *12S* primer set was an appropriate choice for metabarcoding here.

Each metabarcoding extraction was subsampled for three PCR reactions using the MiFish *12S* primer set. PCR amplification for the MiFish primer set was conducted following the thermocycler profile of Curd et al. 2019. MiFish PCR reactions had 25 μL reaction volume containing 12.5 μL QIAGEN Multiplex Taq PCR 2x Master Mix (Qiagen Inc., Valencia, CA, USA), 6.5 µL of molecular grade water, 2.5 µL of each primer (2 µmol/L), and 1 μL DNA extraction. MiFish 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 of a 30s denaturation at 94°C, a 30s annealing that started at 69.5°C and then decreased by 1.5°C for each subsequent cycle (last cycle was 50°C), finishing with a 1 min extension at 72°C. This initial touchdown profile was followed by 35 additional cycles using identical parameters except a constant annealing temperature of 50°C and ending with a final extension at 72°C for 10 min.

Two non-native non-marine vertebrates, American alligator (*Alligator mississippiensis*) and dromedary camel (*Camelus dromedarius*), were purchased at a local market and used as positive controls. For all positive controls, tissues were extracted using the Qiagen Blood and Tissue kit following the manufacturer’s instructions. All PCR products were visualized via electrophoresis on 2% agarose gels to ensure amplification success and correct product size. Only filters from four jars failed to amplify, and upon further inspection within the archived notes, all these samples had known preservation issues (e.g., preservative dried out, observed mold, etc.). All other DNA extractions successfully amplified.

We prepared libraries following the methods of Curd et al. using a two-step PCR amplification method with one final pool per primer set. Previous work indicated that two-step PCR amplification can reduce amplification biases (Gohl et al., 2016; O’donnell et al., 2016)(89, 90) perhaps introduced by the inclusion of various indices during one-step PCR procedures. Variations in the relative amplification efficiency of each PCR is a concern here given the desire to study an array of targets in an oceanic region over space and time. Overall, there are review papers available that outline the advantages and disadvantages for one-step and two-step PCR protocols (Bohmann et al., 2022).

Prior to the second indexing PCR reaction, PCR samples from the first reaction were cleaned using the Serapure magnetic bead protocol. We quantified bead-cleaned samples with the Quant-iT™ broad range dsDNA Assay Kit (Thermofisher Scientific, Waltham, MA, USA) on a Victor3 plate reader (Perkin Elmer Waltham, MA, USA). We indexed the sample libraries using unique combinations of the Nextera Index A, B, C, and D Kit (Illumina, San Diego, CA, USA) and KAPA HiFi HotStart Ready Mix (Kapa Biosystems, Sigma Aldrich, St. Louis, MO, USA). Indexing was performed with a second PCR using a 25 μL reaction mixture containing 12.5 μL of Kapa HiFi Hotstart Ready mix, 1.25 μL of index primers, 10 ng of template DNA to ensure equal copy number, and the remaining volume was filled using molecular grade water depending on cleaned PCR product concentration. Index thermocycling parameters were: 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. To confirm successful PCR and correct product size, we electrophoresed PCR products on 2% agarose gels. We then bead cleaned and quantified DNA concentration, as described above so that we could pool samples so as to have equal copy number for each unique library. Pooled libraries were sequenced on an Illumina NextSeq PE 2x150 at UCLA Technology Center for Genomics and Bioinformatics.

*Bioinformatics*

The resulting metabarcoding data were processed using the Anacapa Toolkit to conduct quality control, amplicon sequence variant (ASV) parsing, and taxonomic assignment using user-generated custom reference databases. We processed sequences using default parameters except using a Q score cutoff of 30 and assigned taxonomy using CRUX-generated metabarcode specific reference databases (Gold et al., 2021). The MiFish sequencing data was assigned taxonomy using the California fish specific reference database and a bootstrap confidence cutoff score of 60 following Gold et al., (2021).

The two resulting raw ASV community tables were decontaminated following Kelly et al., (2018). First, only merged paired reads that occurred at least twice (e.g., no singletons) were retained. Second, we estimated index hopping between samples by calculating the proportion of sequences within the positive control samples and then subtracting reads from each sample by the sample read depth multiplied by the proportion of reads observed in the positive controls. Third, we discarded technical replicates with fewer than 30,000 reads. Fourth, we calculated Bray-Curtis dissimilarities between technical PCR replicates and fit a skewed beta distribution (a= 0.6, b= 9.5). We then removed all replicates with greater than 95% probability of belonging to the beta distribution. Resulting tables were then combined into a final ASV community table in R.

*Microscopy Identification of Ichthyoplankton*

Plankton samples were processed at the NOAA Southwest Fisheries Science Center ichthyoplankton laboratory. From each plankton sample, fish larvae were sorted and identified through microscopy to the lowest practical taxon (McClatchie, 2014; Thompson et al., 2017). Most taxa were identified to species although some were only characterized to genus or family level (See larval\_counts\_20210305.csv). The number of larvae per species per jar, total abundance of filtered ichthyoplankton, and proportion of jar sorted were recorded.

*Estimating Abundance*

We estimated the abundance of ichthyoplankton in each jar using a novel joint Bayesian hierarchical model based on the quantitative metabarcoding framework described in Shelton et al. 2022, detailed here in the supplemental methods. Importantly, we model taxon sequence-read counts from metabarcoding to account for the PCR process in which each taxon is subject to a different amplification efficiency based on the primer set used. We estimate that the number of sequenced amplicons, for any species *i,* is a nonlinear function of the species-specific fraction of DNA in the template (Kelly et al., 2019; McLaren et al., 2019; Silverman et al., 2021 ; we use *i* to represent species, but can be generalized to represent ASVs or other molecular targets). The amplicons produced during a PCR reaction are dictated by the amplicon efficiency parameter , which is characteristic of the interaction between the particular PCR reaction and each species being amplified. Thus, for any species *i*, the number of amplicons should be directly related to the efficiency of amplification and the starting concentration of DNA template such that

(1)

where is amplicon abundance, is the true number of DNA copies in the reaction attributable to species *i*, is the species-specific amplification efficiency (bounded on (0,1)), and is the number of PCR cycles used in the reaction (Lalam, 2006). We note this model assumes that PCR amplification has not approached saturation and therefore the PCR is still amplifying exponentially. We, and others (McLaren et al., 2019), argue this assumption is valid because 1) the total concentration of DNA within a filtered ethanol sample is low (<1 ng/μL), 2) the PCR reagents are supplied in excess and therefore are unlikely to be saturating the PCR, and 3) evidence from previous studies supports these assumptions (McLaren et al., 2019; Shelton et al., 2022; Silverman et al., 2021). However, future models could be developed to account for a saturating PCR curve (Lalam, 2006).

If amplicons could be perfectly observed, Equation 1 would faithfully relate amplicon abundance to the biological value of interest, , the true number of template DNA copies. Unfortunately, standard metabarcoding does not allow for such direct observation of amplicon abundance because, unlike in qPCR amplification of a single target, the production of all the varieties of amplicons generated during a sequencing run cannot be tracked, and are not amenable to simple quantification due to combined effects of the PCR process and subsampling.

To illustrate this point, the number of amplicons expected for any species with > 0 is very large due to being a large number and typically not being close to 0, (e.g. with = 2, = 0.75, and = 36, = 1.12 × 109). Thus, given there are typically many species being amplified simultaneously, a single reaction can produce 1010 or more DNA copies with the actual number driven primarily by the values among species and . Importantly, not all molecules of DNA are transferred through each molecular step, particularly as DNA sequencing machines do not read all of the copies from such a reaction; they read only a small fraction of the reads (on the order of 106 to 107 reads) (Egozcue et al., 2020; Silverman et al., 2020). This subsampling changes what in Equation 1 appears to be a single-species process – each species being amplified independently – into a multi-species process where the number of amplicons observed for species *i* depends upon both the amplicons produced for species *i* = 1 and the amplicons produced for species *i* = 2, 3, ..., *I* in the same reaction. Observations of amplicons are thus compositional data, meaning they are the proportions of the sample amplicon reads and therefore convey relative quantitative information of the observed species, and therefore need to be analyzed as such (Gloor et al., 2017).

To harness the ability to generate quantitative data from Equation 1 as much as possible, we develop a model for a single sample with many species. As above, if we let *I* index species with *I* = 1, 2, ..., *I*, then we can write a deterministic Equation for the number of amplicons observed in log-space as

(2)

where the only new term is η, representing the proportion of reads observed from a given sampling run. Note that in this formulation η is a single value shared across all species in a sample and serves to scale the number of amplicons observed. Additionally, we can rewrite the number of DNA copies in terms of the proportional number of larvae counts, such that log() = log() − log(). Note that the second term in this equation, log(), is a sum of the counts across all species, and so is a single shared value for all species. As such it can be integrated into the value η that scales the overall abundance for each species *i*,

(3)

This equation is appealing because it provides a process-oriented description of the biology of interest (the β parameters), a term for how PCR modifies our amplicon sequence count observation of the true abundance ( , and a term for how subsampling of DNA reads will adjust the number of amplicons observed (log(η)). This third term also links all of the single-species components to produce a multi-species model. It is important to note that while both Equations 2 and 3 use the term η, the interpretation and plausible range of this parameter are distinct in the two equations. In Equation 2, 0 < η ≤ 1, while in Equation 3 η is not constrained to be less than 1 (η > 0).

In practice, PCR and subsampling are not deterministic but random processes (Egozcue et al., 2020). Furthermore, we are rarely interested in results from a single sample but rather multiple samples collected across sites *j* and times *t*. In addition, we let λ*ijtk* be the expected number of amplicons observed, with *k* indexing unique PCR reactions to account for the fact that there may be multiple individual PCR reactions for a single collected sample,

(4)

In this case, is assumed to be constant for each species among all sites, times, and PCR reactions (this assumption is strongly supported by McLaren et al., 2019; Shelton et al., 2022; Silverman et al., 2021). We incorporate stochasticity by allowing to the number of observed amplicons to vary from the deterministic mean by specifying the observed values as following a negative binomial distribution,

(5)

(6)

where the expected value and variance of are and , respectively. Note that we allow for the scale parameter to vary with the predicted mean, such that the amount of dispersion in the negative binomial shifts to be large when λ is small and to decrease as λ increases.

However, by itself, this model has substantial identifiability problems; in the absence of additional information, it is not possible to estimate the β and *a* parameters from metabarcoding data alone. Including morphological count data enables us to estimate the confounded parameters by bounding additional information about the underlying species abundances. Below we discuss how these two datasets are integrated (see Shelton et al. (2022) for the application of mock community data to similarly calibrate metabarcoding data).

For each sampled station, we have two independent sets of observed data: 1) counts of larval/juvenile fishes for each taxon from the formaldehyde jars (; indexes as above) and 2) counts of amplicons for each taxon from ethanol jars (). These observed data arise from a common (but unobserved) biomass for each species at each station-year combination ( a latent (unobserved) variable).

We assume that the data are counts for each species in each sample, , derived from the true density of each species , the fraction of total specimens counted in each vial, , and the volume of water filtered for that station relative to a standard volume, ; ≈ 1 for most samples, < 1 indicates a smaller volume of water was sampled.

(7)

+ (8)

We consider to be the true proportion of biomass at a given station-year for each taxon *i*, .

We note that microscopy counts were modeled as Poisson-distributed given their relatively small absolute values and low variance (Thompson et al., 2017), and amplicon sequence data were modeled using a Negative Binomial distribution given their relatively high absolute values and high variability among replicates (Figure 1). These statistical distributions are commonly used in models of count and amplicon data, respectively (Chambert et al., 2018; Meyer-Gutbrod et al., 2021; Ren & Kuan, 2020).

To combine our information from the manual counts and metabarcoding, we need to recognize that our observations ( and ) are linked together by a common variable () and thus we can model them jointly (Hobbs & Hooten, 2015). We represent the amplification process using Equations 5 and 6 above (amplicons were sequenced in triplicate reactions for each jar). The manual counts are modeled as in Equations 7 and 8.

Our model assumes the fraction of template DNA in each sample is proportional to the counts of individual species’ larvae in each paired jar (McLaren et al., 2019). Moreover, we assume that in each sample there is template DNA from species that are uncounted, unidentifiable, or otherwise unobserved (Egozcue et al., 2020). In practice, this DNA might derive from stochastic sampling between each side of the bongo net, excreted waste, stray tissue, cells, or microscopic genetic material extracted along with the visible larvae.

The above is sufficient if all of the species identified by morphological counts are identical to the species identified by the genetic methods. But this is often not the case; some larvae are not separable to species based on morphology and some species are not separable to species based on a single genetic primer. Furthermore, some species do not amplify at all in the PCR (≈ 0) or else are undetected, being swamped out by the far-more-common amplicons of other species. To accommodate non-overlapping sets of species among sampling methods we introduce a new variable, , which specifies the true (*M* is for “main”) density of species *i* at site *j* and time *t*. We assume that there is a mapping between this main density and the density observed by each sampling method. Specifically, we assume the species in the main list maps uniquely on to no more than one taxonomic group in each observation method, but multiple main species can map onto a single group for each observation method. For example, if the observation of larval counts identified a specimen as *Sebastes* sp., we assume this may map onto one or more specific taxa (e.g., *Sebastes paucispinis*) in the main list, but conversely, *Sebastes paucispinis* on the main list may not map to more than one entity identified by each observation method.

We can construct a mapping matrix,, that transforms the species in the main list, (a vector of length , the number of true species in the sample) into the species grouping observed by sampling method *S*, (a vector of length , the number of groups observed by method *S*). We drop the *j* and *t* subscript because this mapping does not depend on the identity of the community being sampled. Then,

(9)

is a by matrix.

For example, if there are four species in the community and methodonly observes three groups, the matrix could look like this

(10)

This might happen if species 2 and species 4 (columns 2 and 4, respectively) were from the same genus and the PCR primer from method *S* can only resolve those two species at the genus level. To provide a further example, take an invented community of four species with individuals in the community. The true community as observed through method *S* would be

(11)

and so is a linear combination of the true community. Of course, there is no requirement that elements of be integers, but that makes the above example easy and transparent.

It is easy to incorporate this added complexity into the models in the previous section. If we assign method *S* to be manual counts and *W* to be the Mifish PCR primer, we need to construct a main list of species to define and build two mapping matrices, and, that determine which species or species-groups are observed by each method. We can then add a subscript for each additional method and use the same form as above. For example,

+ (12)

(13)

And with additional sampling methods, we can make different mappings from the true abundance to the observations of each method.

We develop and fit the above model in a Bayesian framework using the Stan language, as implemented in *RStan* (Stan Development Team, 2021). All code is available as supplementary material. Table S2 provides prior distributions used in the model.

We ran five MCMC chains with 1,000 warmups and 4,000 sampling iterations. We retained every other MCMC sample. We initiated each chain at randomly determined starting values. The model converged ( < 1.01; Gelman–Rubin diagnostics) and had no divergent transitions. We performed standard posterior predictive checks to assess model fit.

**Discussion**

Previous work has suggested that PCR reactions are not perfectly exponential but saturate over time as reagents are consumed and enzymatic activity declines in efficiency (Boggy & Woolf, 2010; Chatterjee et al., 2012; Kubista et al., 2006). If the employed PCR reaction were to indeed follow a saturating function, then the number of successful PCR cycles would be effectively lower than the full 49 cycles employed here. Thus, the resulting amplification efficiency coefficients generated by our model would be underestimated. However, given the function of the PCR equation, such a bias would be applied uniformly as a scalar across all estimated amplification efficiencies and thus not affect the broad interpretation of our model results. Fortunately, accounting for any impact of PCR saturation can be easily remedied with the incorporation of a decay coefficient in the function of the underlying PCR equation in future modeling efforts (Boggy & Woolf, 2010; Chatterjee et al., 2012; Kubista et al., 2006). Future work understanding the fundamental mechanisms of PCR amplification and bias, particularly whether PCR follows an exponential or saturating function, is warranted given the importance of metabarcoding approaches to biological observations (Gold et al., 2022; M. R. McLaren et al., 2019; Shelton et al., 2022; Silverman et al., 2021).

We also assumed that the observed morphological counts directly correlate with the starting DNA concentration in each jar. However, this assumption is likely invalid for a few reasons. First, our work did not incorporate morphological counts of egg abundance from the jars. Almost undoubtedly, our DNA metabarcoding approaches successfully detected DNA derived from eggs as well as any damaged larvae that could not be morphologically identified within the samples. Thus, a shortcoming of our model used here is the underestimation of the true concentration of DNA within the sampled preservative when morphological identifications are not possible, leading to overestimates of amplification efficiencies and underestimates of ichthyoplankton abundances (Figure 3b,c).

In addition, all morphological counts and eDNA data were conducted in paired bongo net tows which have subtle differences in ichthyoplankton abundance and composition, particularly for low abundant species (Thompson, McClatchie, et al., 2017). Thus, the stochastic nature of detecting rare ichthyoplankton across both cod ends may help explain the higher observed rates of non-detection across technical PCR replicates for rare (n<9 individuals) fish larvae.

Furthermore, we know that both metabarcoding and morphological taxonomic identification are imperfect, leading to mismatches between survey approaches that result in undercounting of larvae or DNA for specific taxa. Unsurprisingly, many taxa lack morphological characteristics, especially at early developmental life stages (H. G. Moser, 1996). Likewise, many taxa lack sufficient genetic differentiation to resolve species identification either because reference databases are incomplete despite extensive efforts (Gold et al., 2021) or because many adaptive radiations like the genus *Sebastes* have identical sequences for the employed marker sets (Min et al., 2021). These limitations for genetic identification can be overcome with additional markers and more comprehensive reference databases (Schenekar et al., 2020). Although we took great efforts to develop a species mapping index that linked observations between both methods, all survey approaches will be limited by imperfect taxonomic identification between both sampling techniques.

A combination of the above limitations led to the observed non-detections in morphological data sets. For example, in one sample from 2014 nearly 90% of sequence reads were identified to *Citharichthys sordidus*, despite morphological methods failing to detect this species within the jar. Given that larvae from this species are regularly identified in the CalCOFI program, this non-detection was likely driven by a high occurrence of eggs, which are not morphologically identifiable from other congeners in the Family Pleuronectidae. The high occurrence of eggs within a sample would result in a non-detection in the morphological data set while metabarcoding data would be swamped by *Citharichthys sordidus* DNA. Thus we hypothesize that patterns of high DNA abundance associated with non-detections in the morphological data sets are driven by our exclusion of eggs and other sources of DNA in the morphological data set.

**Table S1: MiFish Universal Primers Used.**

The MiFish Universal primer used here is bolded and underlined. The Illumina Nextera adapter.

|  |  |
| --- | --- |
| MiFish Universal F With Illumina Nextera Adapter | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG**GCCGGTAAAACTCGTGCCAGC** |
| MiFish Universal R With Illumina Nextera Adapter | GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG**CATAGTGGGGTATCTAATCCCAGTTTG** |
| MiFish Universal F | **GCCGGTAAAACTCGTGCCAGC** |
| MiFish Universal R | **CATAGTGGGGTATCTAATCCCAGTTTG** |

**Table S2: Prior and parameter descriptions for the Stan Model.**



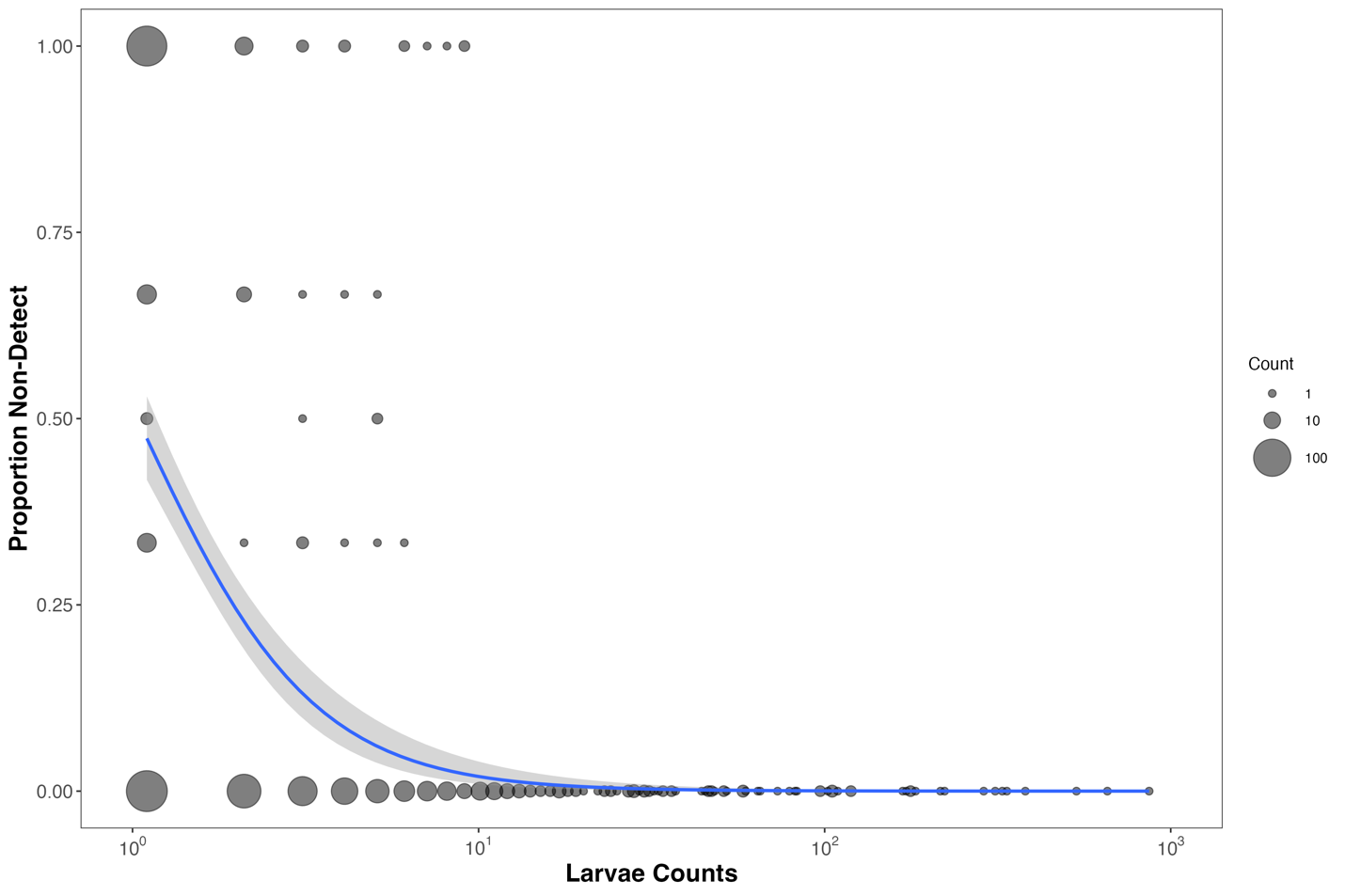
Supplemental Figures

**Chart, scatter chart

Description automatically generated**

**Figure S1. Proportional Abundance Comparisons Between *Stenobrachius leucopsarus variants***

Find poor correlation between the proportional abundance (sequence reads/total reads per sample) of the two *Stenobrachius leucopsarus* variants. The one to one line is plotted in red.



**Figure S2. Non-detection in Metabarcoding Data are Strongly Correlated with Larvae Counts**

Non-detections (zeros across multiple technical PCR replicates) in metabarcoding data are strongly correlated with larvae counts. Proportion of non-detections are calculated per set of technical replicates per DNA extraction per species. A value of 1 corresponds to non-detection across all technical replicates while a value of 0 indicates it was detected in all technical replicates. The size of the points are proportional to the count of non-detection occurrences for a given number of larvae. We did not observe any non-detections when larvae counts were greater than 9 in a jar. These results suggest a limit of detection for metabarcoding results.

Diagram

Description automatically generated

**Figure S3. Novel Marine Heatwave Assemblages Derived from Only Morphological Data**

Morphologically derived shifts in species abundances with the onset of the marine heatwave (1996–2013 vs. 2014–2019). Synchronous increases in southern mesopelagic species and Northern Anchovy (*Engraulis mordax*) were observed across all stations. Stations are in rows, species in columns, and the joint model estimated change in abundance between the two ecological phases is shown as the response variable. Fisheries targets including Pacific Sardine (*Sardinops sagax*) and North Pacific Hake (*Merluccius productus*), as well as many other benthic and coastal species, had concurrent negative associations. Significant differences during and after the marine heatwave are marked with + or -.

Graphical user interface

Description automatically generated with medium confidence

**Figure S4. Novel Marine Heatwave Assemblages Derived from Only Morphological Data Excluding Abundant Species**

Morphologically derived shifts in species abundances with the onset of the marine heatwave (1996–2013 vs. 2014–2019). This plot is identical to S4 except the three most abundant taxa (*Engraulis mordax, Sebastes sp.,* and *Leuroglossus stilbius*) were excluded. Stations are in rows, species in columns, and the joint model estimated change in abundance between the two ecological phases is shown as the response variable. Fisheries targets including Pacific Sardine (*Sardinops sagax*) and North Pacific Hake (*Merluccius productus*), as well as many other benthic and coastal species, had concurrent negative associations. Significant differences during and after the marine heatwave are marked with + or -.

Graphical user interface, chart, application, Word, Excel

Description automatically generated

**Figure S5. Synchronous Increase in Morphologically Derived Anchovy Abundance During and After Marine Heatwave**

Morphologically derived larval fish abundances (larvae counts per standardized volume towed) over time at each of the four sampled stations. Morphological counts documented increases in Northern Anchovy (*Engraulis mordax*) [blue] during the recent Pacific Marine Heatwave and low spawning of Pacific Sardine (*Sardinops* sagax) [red] over the past decade (points are total counts; the shaded region is during and after the marine heatwave).